United States
Environmental Protection
Agency

Office of Water
4606

EPA 816-R-99-006
November 1999

A Review of Contaminant Occurrence in Public Water Systems

CONTENTS

TABLES AND FIGURES

ACRONYMS

Advanced Notice of Proposed Rule Making (ANPRM)

Alternative Monitoring Guidelines (AMG)

American Water Works Service Company (AWWSC)

Chemical Monitoring Reform (CMR)

Comma Separated Values (CSV)

Community Water System (CWS)

Dense Non-aqueous Phase Liquids (DNAPLS)

Dibromochloropropane (DBCP)

Environmental Protection Agency (EPA)

Ethylene Dibromide (EDB)

High-occurrence Synthetic Organic Chemical (HiSOC)

Inorganic Chemical (IOC)

Light Non-Aqueous Phase Liquids (LNAPLs)

Low-occurrence Synthetic Organic Chemical (LoSOC)

Maximum Contaminant Level (MCL)

Method Detection Limit (MDL)

Minimum Reporting Level (or Limit, MRL)

National Primary Drinking Water Regulations (NPDWRs)

National Water Quality Assessment Program (NAWQA)

National Alachlor Well Water Survey (NAWWS)

National Pesticide Survey (NPS)

Non-Transient Non-Community Water System (NTNCWS)

North-central (NC)

Northeastern (NE)

Office of Ground Water and Drinking Water (OGWDW)

Percentage of Systems with Exceedances (>MCL)

Percentage of Systems with Detections (>MRL)

Permanent Monitoring Relief (PMR)

Public Water System (PWS)

Public Water System Identifier (PWSID)

Regulatory Implementation Branch (RIB)

Safe Drinking Water Act (SDWA)

Safe Drinking Water Information System (SDWIS)

Source Water Assessment Program (SWAP)

Synthetic Organic Chemical (SOC)

Tetrachloroethylene (PCE)

Toxic Release Inventory (TRI)

Transient Non-Community Water System (TNCWS)

Tricloroethylene (TCE)

United States Geological Survey (USGS)

Unregulated Contaminant Monitoring Information System (URCIS)

Volatile Organic Chemical (VOC)

USEPA, Office of Water Report: EPA 816-R-99-006, November 1999

A REVIEW OF CONTAMINANT OCCURRENCE IN PUBLIC DRINKING WATER SYSTEMS:

EXECUTIVE SUMMARY

This study represents the Environmental Protection Agency’s (EPA’s) most extensive analysis to date of the occurrence of contaminants in Public Water Systems (PWSs) regulated under the Safe Drinking Water Act (SDWA). EPA's Office of Ground Water and Drinking Water conducted this study to provide a better scientific basis for considering changes to the chemical monitoring requirements in the drinking water program.

State databases, comprised of SDWA compliance-monitoring data from PWSs, were the primary data sources for this analysis. Data from 12 states were analyzed in detail. Data from 8 States were used to develop a national cross-section of contaminant occurrence. The States were selected to represent the national range of hydrologic regimes and pollution potential. The data used represent more than 10.7 million analytical results from nearly 26,000 PWSs. More than 70% of the data are from 1993 or later; reflecting the beginning of the Phase II/V requirements.

While the national cross-section cannot be stated to be “statistically representative,” the results are clearly indicative of the national values for occurrence. The eight cross-section States represent over 25% of the population using public water supplies, as well. Comparisons with other limited national data suggest that the cross-section occurrence values are slightly high yet very comparable to other national averages.

Additional sources of national data were also used, such as: EPA's Unregulated Contaminant Monitoring Information System (URCIS), which includes data from 40 States and territories from the first round of unregulated contaminant monitoring; and U.S. Geological Survey studies of ambient monitoring, including results from the National Water Quality Assessment Program, covering parts of 42 States.

Data management and handling were an important component of this effort because each State database required unique editing and formatting to facilitate consistent analysis and valid comparisons among States. Variables were cross-checked for consistency and unique issues were resolved in consultation with the States. Whenever errors or ambiguities could not be resolved, data were eliminated to avoid aberrant results. This analysis summarizes contaminant occurrence by systems, to avoid the skew that is inherent in summaries by sample. Only standard SDWA compliance samples were used.

Contaminant occurrence in drinking water systems is highly variable. All 64 Phase II/V contaminants have been detected in drinking water systems, however, the frequency of occurrence in water systems ranges widely. Fifty-nine of the 64 contaminants have a reported detection at greater than half their Maximum Contaminant Level (MCL); but MCL exceedances are not common. The five that do not have detections >0.5 MCL are all Synthetic Organic Chemicals (SOCs). Fifty-five of the contaminants occur in more than 1% of PWSs. The nine that occur in less than 1% of PWSs are also all SOCs: alachlor, carbofuran, chlordane, glyphosate, hexachlorobenzene, hexachloropentadiene, oxamyl, toxaphene, and PCBs (though PCBs have not been monitored as intensively as other SOCs).

Twenty-nine of the 30 regulated SOCs have not been detected at all in some States (only phthalate occurs in every State, but its occurrence relates, in part, to sample contamination). The greatest range for any of the organic chemicals is for atrazine, which ranges from no systems with detections in some States to 97% of surface water systems showing detections in Midwestern States.

Seven of the 21 Volatile Organic Chemicals (VOCs) occur in every State studied: ethylbenzene, cis-1,2- dichloroethane, tetrachloroethylene (PCE), trichloroethylene (TCE), vinyl chloride, 1,1,1- trichlorethane, and xylenes. It is common for many VOCs to occur in about 30 percent of surface or ground water systems.

Surface water systems tend to be more vulnerable than ground water systems for most contaminants. SOCs are far more common in surface water systems than ground water systems and surface water systems have more MCL exceedances as well. VOCs are more common than most SOCs in both surface and ground water systems. VOCs show greater general occurrence in surface water but surface water systems and ground water systems are nearly equal for the occurrence of exceedances. Ground water systems show slightly more MCL violations for VOCs. Inorganic Chemicals (IOCs) are about equally common in surface water and ground water systems, but ground water systems generally have higher average concentrations. IOCs have relatively high occurrence rates because they occur naturally.

For most VOCs and many SOCs, larger systems show proportionately greater contaminant occurrence than smaller systems, and a tendency, at least for VOCs, to have proportionately more systems with exceedances. This trend is most consistent for ground water systems. To get the greater ground water yields needed, larger systems typically must use unconfined ground water supplies, which are typically shallow and more vulnerable to contaminant occurrence. While large systems have proportionately greater occurrence the number of small systems is much greater. Hence, there is a large number of small systems affected by contaminants.

Ground water systems are complex and it is difficult to make generalizations about vulnerability. Information on aquifer characteristics, well depth, well casing and construction details, wellhead protection information, as well as system and well maintenance history are some of the factors that might be considered to establish specific monitoring approaches for a ground water system.

Many SOCs (pesticides, in particular) exhibit strong seasonal patterns because their application or discharge into the environment is concentrated seasonally. In contrast, VOCs do not typically show such seasonality. Targeting monitoring to such vulnerable periods can improve the effectiveness of compliance monitoring and the accuracy of exposure estimates.

Statistical reviews suggest that sampling strategies can be designed to account for seasonal peak occurrence. While vulnerable periods may vary from region to region, the data suggest that only one vulnerable period would likely need to be targeted in a given region. Vulnerable periods are apparent for surface waters and shallow, vulnerable ground water systems. For deeper ground water systems few generalizations can be made, except to note that as well depth increases the degree of temporal variability usually decreases.

ACKNOWLEDGMENTS

The compilation and analysis of data presented in this report was undertaken by EPA's Office of Ground Water and Drinking Water (OGWDW) to enhance the scientific understanding of the occurrence of chemical contaminants in public drinking water systems and to refine the basis for the monitoring of these contaminants. This effort was directed by Mr. Edwin Thomas of OGWDW. It began under the management of Mr. Michael Muse.

We would like to thank the many States, as well as the American Water Works Service Company that contributed data sets and valuable advice. Thanks also to the many public water systems that conducted the monitoring. EPA's Office of Pesticide Programs and the U.S. Geological Survey contributed valuable supplemental studies and data. Mr. Lewis Summers and Mr. Charles Job of OGWDW managed the completion of EPA's URCIS data base. The report received extensive review from EPA and State personnel, as well as external stakeholders and expert peer reviewers. In particular, thanks are due Robert Libra (Geological Survey Bureau, Iowa Department of Natural Resources), Robert Gilliom and John Zogorski (Water Resources Division, National Water Quality Assessment, US Geological Survey), and Kenneth Cantor (Occupational Epidemiology Branch, National Cancer Institute) for their constructive reviews and suggestions.

The Cadmus Group, Inc. served as the prime contractor for this project, supporting the data analysis and report development. Dr. George Hallberg served as Cadmus' Project Manager. For further information, contact Ed Thomas, USEPA, OGWDW, 202-260-0910 (fax, 202-401-2345) or e-mail, thomas.edwin@epa.gov.

DISCLAIMER

This report does not constitute U.S. Environmental Protection Agency Policy. Mention of trade names or commercial products does not constitute endorsement or recommendation for use.

This document is designed to provide technical background for the Office of Ground Water and Drinking Water's program. The document does not, however, substitute for the Safe Drinking Water Act or EPA's regulations nor is this document a regulation itself. Thus, it cannot impose legally-binding requirements on EPA, States, or the regulated community, and may not apply to a particular situation based on the circumstances.

A REVIEW OF CONTAMINANT OCCURRENCE IN PUBLIC DRINKING WATER SYSTEMS

I. INTRODUCTION

The purpose of this report is to review the analysis of data on the occurrence of chemical contaminants in public drinking-water systems regulated under the Safe Drinking Water Act (SDWA). This is the most extensive analysis EPA has conducted to date for these contaminants and it provides an overview of several occurrence issues that can contribute improved scientific understanding for management of SDWA programs.

I.A. Background

EPA's Office of Ground Water and Drinking Water (OGWDW) is required by the Safe Drinking Water Act (SDWA) to protect public health and safety by setting standards for the quality of water provided by public water systems. To meet the requirements of the SDWA, EPA has published the National Primary Drinking Water Regulations (NPDWRs) to establish Federal standards for drinking water contaminants and rules for monitoring those contaminants. EPA promulgated the first regulations, for nine inorganic chemicals, six organic chemicals, total coliform bacteria, and turbidity, in 1975. Subsequently, EPA has published numerous amendments and standards for additional contaminants.

The Phase rules regulate 64 chemical contaminants (Table I.A.1) that are of primary concern because of their potential to cause chronic health effects and their known occurrence in the source waters of public water systems. The NPDWRs for Phase I were promulgated in July 1987. The Phase II regulations, which modified the requirements for Phase I contaminants, were promulgated in 1991. Standards for additional contaminants were established with the Phase IIB regulation of 1991 and the Phase V regulation of July 1992. The Phase rules have been challenged by some as overly prescriptive, complicated, and costly. Critics have contended that nationally uniform rules often result in requirements for monitoring contaminants that do not occur in many different regions of the country. However, there has been limited data or analysis of data to support these criticism or to inform deliberations for reform.

The Regulatory Implementation Branch (RIB) (formerly the Drinking Water Branch) of the OGWDW is responsible for implementing these regulations and recommending revisions to improve the efficiency of implementation and reduce the regulatory burden. In 1995, the RIB began a review to assess ways to simplify requirements and facilitate implementation of its many programs under SDWA. Simultaneously, OGWDW was responding to a Presidential directive to improve all Agency programs and reduce regulatory burden. Streamlining the monitoring requirements of the Phase II, IIB, and V rules, described as Chemical Monitoring Reform (CMR), offers an opportunity to respond to both initiatives.

RIB established the CMR State Workgroup to develop a framework for proposed CMR changes to the Phase rules. During the deliberations to develop CMR, Congress passed the 1996 Amendments to SDWA. In these Amendments, Congress authorized States that have received

EPA approval of their Source Water Assessment Programs (SWAP) to offer Permanent Monitoring Relief (PMR) to public water systems. The provisions of the PMR have been considered under the Alternative Monitoring Guidelines (AMG), and most components have been incorporated into the CMR options that EPA is considering. The 1996 Amendments also directed EPA to review the monitoring requirements for not fewer than 12 contaminants identified by the Administrator and to promulgate any necessary modifications by August 6, 1998. These actions altered the time line for development and proposal of the CMR.

In June 1997, an Advanced Notice of Proposed Rule Making (ANPRM) was issued for CMR and AMG. The ANPRM described some options under consideration and asked for additional public and stakeholder input. EPA's limited analysis of occurrence data, developed for the ANPRM, as well as the comments and response to the ANPRM and the CMR State Workgroup, all indicated the need for a more comprehensive review of contaminant occurrence and assessment of the implications of these data for designing chemical monitoring reform options. The CMR revisions were delayed until occurrence data could be further evaluated. On August 5, 1997, the Agency announced the final guidelines for AMG.

A preliminary review of occurrence data was presented at a CMR Stakeholders' Meeting in Spring 1998. The Stakeholders also concurred that further action on CMR should be postponed until the data could be analyzed more fully. In July 1998, a Federal Register Notice announced that no further action was warranted or would be taken on CMR until the occurrence data had been analyzed and the results evaluated.

This report summarizes the results of extensive additional analysis of contaminant occurrence data and reviews these findings to inform continuing deliberations for possible changes in chemical monitoring strategies.

Table I.A.1.a. Inorganic chemicals (13) regulated in PWSs under SDWA Phase rules.

	Contaminant	**MCL mg/L**	**Common Sources of Contaminant in Drinking Water**	**Potential Health Effects**
Inorganic Chemicals				
	Antimony (total)	0.0060	Fire retardants, ceramics, electronics, fireworks, solder	Cancer
	Arsenic	0.050	Natural deposits, smelters, glass, electronic wastes, old orchards	Skin, nervous system toxicity
	Asbestos	7 mf/L	Natural deposits; asbestos cement in water systems	Cancer
	Barium	2.0	Natural deposits; pigments, epoxy sealants, spent coal	Circulatory system effects
	Beryllium (total)	0.0040	Electrical, aerospace, defense industries	Bone, lung damage
	Cadmium	0.0050	Galvanized pipe corrosion; natural deposits; batteries, paints	Kidney effects
	Chromium	0.10	Natural deposits; mining, electroplating, pigments	Liver, kidney, circulatory disorders
	Cyanide	0.20	Electroplating, steel, plastics, mining, fertilizer	Thyroid, nervous system damage
	Fluoride	4.0	Natural deposits; fertilizer, aluminum industry; H2O additive	Skeletal and dental fluorosis
	Mercury	0.0020	Crop runoff; natural deposits; batteries, electrical switches	Kidney, nervous system disorders
	Nickel	0.10	Metal alloys, electroplating, batteries, chemical production	Heart, liver damage
	Selenium	0.050	Natural deposits; mining, smelting, coal/oil combustion	Liver damage
	Thallium (total)	0.0020	Electronics, drugs, alloys, glass	Kidney, liver, brain, intestinal disorders

Table I.A.1.b. Synthetic organic chemicals (30) regulated in PWSs under SDWA Phase rules.

	Contaminant	MCL mg/L	Common Sources of Contaminant in Drinking Water	Potential Health Effects
Synthetic Organic Chemicals				
	Alachlor (Lasso)	0.0020	Herbicide on corn, soybeans, other crops	Cancer
	Atrazine	0.0030	Herbicide on corn and non-cropland	Mammary gland tumors
	Benzo[a]pyrene	0.0002	Coal tar coating; burning organic matter; volcanoes, fossil fuels	Cancer
	bis(2-ethylhexyl) Adipate	0.40	Synthetic rubber, food packaging, cosmetics	Decreased body weight
	bis(2-ethylhexyl) Phthalate	0.0060	PVC and other plastics	Cancer
	Carbofuran (Furadan)	0.040	Soil fumigant on corn and cotton; restricted in some areas	Nervous, reproductive system disorders
	Chlordane	0.0020	Leaching from soil treatment for termites	Cancer
	2,4-D	0.070	Herbicide on wheat, corn, rangelands, lawns	Liver and kidney damage
	Dalapon	0.20	Herbicide on orchards, beans, coffee, lawns, road/railways	Liver, kidney disorders
	Dibromochloropropane (1,2-Dibromo-3-chloropropane; or DBCP)	0.0002	Soil fumigant on soybeans, cotton, pineapple, orchards	Cancer
	Dinoseb	0.0070	Herbicide on crop and non-crop applications	Thyroid, reproductive organ damage
	Diquat	0.020	Herbicide on land & aquatic weeds	Liver, kidney, eye effects
	Endothall	0.10	Herbicide on crops, land/aquatic weeds; rapidly degraded	Liver, kidney, gastrointestinal disorders
	Endrin	0.0020	Pesticide on insects, rodents, birds; restricted since 1980	Liver, kidney, heart damage
	Ethylene dibromide (1,2-Dibromoethane; or EDB)	0.00005	Leaded gas additives; leaching of soil fumigant	Cancer
	Glyphosate (Roundup)	0.70	Herbicide on grasses, weeds, brush	Liver, kidney damage
	Heptachlor	0.00040	Leaching of insecticide for termites, very few crops	Cancer
	Heptachlor Epoxide	0.00020	Biodegradation of heptachlor	Cancer
	Hexachlorobenzene	0.0010	Pesticide production waste by-product	Cancer
	Hexachlorocyclopentadiene	0.050	Pesticide production intermediate	Kidney, stomach damage
	Lindane	0.00020	Insecticide on cattle, lumber, gardens; restricted 1983	Liver, kidney, nerve, immune, circulatory system
	Methoxychlor	0.040	Insecticide for fruits, vegetables, alfalfa, livestock, pets	Growth, liver, kidney, nerve disorders
	Oxamyl (Vydate)	0.20	Insecticide on apples, potatoes, tomatoes	Kidney damage
	PCBs as Decachlorobiphenyl; PCB total as DCBP	0.00050	Coolant oils from electrical transformers; plasticizers	Cancer
	Pentachlorophenol	0.00100	Wood preservatives, herbicide, cooling tower wastes	Cancer, liver and kidney effects
	Picloram (Tordon)	0.50	Herbicide on broadleaf and woody plants	Kidney, liver damage
	Simazine	0.0040	Herbicide on grass sod, some crops, aquatic algae	Cancer
	2,3,7,8-TCDD (Dioxin)	3x10-8	Chemical production by-product; impurity in herbicides	Cancer
	2,4,5-TP (Silvex)	0.050	Herbicide on crops, right-of-way, golf courses; canceled 1983	Liver and kidney damage
	Toxaphene	0.0030	Insecticide on cattle, cotton, soybeans; canceled 1982	Cancer

Table I.A.1.c. Volatile organic chemicals (21) regulated in PWSs under SDWA Phase rules.

	Contaminant	MCL mg/L	Common Sources of Contaminant in Drinking Water	Potential Health Effects
Volatile Organic Chemicals				
	Benzene	0.0050	Some foods; gas, drugs, pesticide, paint, plastic industries	Cancer
	Carbon tetrachloride	0.0050	Solvents and their degradation products	Cancer
	Chlorobenzene	0.10	Waste solvent from metal degreasing processes	Nervous system and liver
	o-Dichlorobenzene	0.60	Paints, engine cleaning compounds, dyes, chemical wastes	Liver, kidney, blood cell damage
	p-Dichlorobenzene	0.0750	Room and water deodorants, and "mothballs"	Cancer
	1,2-Dichloroethane	0.0050	Leaded gas, fumigants, paints	Cancer
	cis-1,2-Dichloroethylene	0.070	Waste industrial extraction solvents	Liver, kidney, nervous, circulatory disorders
	trans-1,2-Dichloroethylene	0.10	Waste industrial extraction solvents	Liver, kidney, nervous, circulatory disorders
	1,1-Dichloroethene (1,1-Dichloroethylene)	0.0070	Plastics, dyes, perfumes, paints	Cancer, liver & kidney effects
	1,2-Dichloropropane	0.0050	Soil fumigant; waste industrial solvents	Liver, kidney effects; cancer
	Ethyl benzene	0.70	Gasoline; insecticides; chemical manufacturing wastes	Liver, kidney, nervous system disorders
	Methylene chloride (Dichloromethane)	0.0050	Paint stripper, metal degreaser, propellant, extraction	Cancer
	Styrene	0.10	Plastics, rubber, resin, drug industries; leachate from city landfills	Liver, nervous system damage
	Tetrachloroethylene	0.0050	Improper disposal of dry cleaning and other solvents	Cancer
	Toluene	1.0	Gasoline additive; manufacturing and solvent operations	Liver, kidney, nervous, circulatory disorders
	1,2,4-Trichlorobenzene	0.070	Herbicide production; dye carrier	Liver, kidney damage
	1,1,1-Trichloroethane	0.2000	Adhesives, aerosols, textiles, paints, inks, metal degreasers	Liver, nervous system effects
	1,1,2-Trichloroethane	0.0050	Solvent in rubber, other organic products; chemical production wastes	Kidney, liver, nervous system
	Trichloroethene (Trichloroethylene)	0.0050	Textiles, adhesives and metal degreasers	Cancer
	Vinyl chloride	0.0020	May leach from PVC pipe; formed by solvent breakdown	Cancer
	Xylenes (Total)	10.0	By-product of gasoline refining; paints, inks, detergents	Liver, kidney; nervous system disorders

II. DATA USED AND METHODS OF ANALYSIS

The data included in the CMR contaminant occurrence analysis came from various sources including State occurrence databases, EPA databases, private water system databases (e.g., American Water Works Service Company), special studies conducted by States or industry, and published data and reports of the U.S. Geological Survey (USGS). These major data sources, which were analyzed in detail, are summarized in Table II. Additional data were analyzed and other findings were reviewed from various States, industry and published research. These additional sources are not included in Table II but will be introduced and referenced where appropriate in this report. These additional databases were not completely analyzed for several reasons. In some cases necessary data elements were incomplete, formatting problems impeded analysis, or their data was of limited applicability. In other cases they were redundant with more complete databases.

The State databases, providing SDWA compliance monitoring data from public drinking water systems, were the most important sources for this analysis. These data form the core of the results presented in this report. As summarized in Table II, data from 12 States were analyzed in detail, providing the summary information described in this report. These data represent more than 10.7 million analytical results from nearly 26,000 public water systems. In most cases, the State databases contained additional data that were not included either because they were not appropriate for this analysis or because they posed various data quality problems (as discussed in later sections). (As noted on Table III, data from 14 States were reviewed.)

As shown in Table II, various States may include data from different time periods. However, the majority of data are from 1993 and later (especially the data used in this summary analysis), which marks the beginning of Phase II/V monitoring. Initial screening of the data showed that most data quality problems were in pre-1993 data. Therefore, in many of the data sets, results gathered before 1993, or 1990 in particular, were eliminated from these analyses. More than 70% of all data utilized in this report are from 1993 or later; this proportion is even greater for most synthetic organic chemicals (SOCs).

These databases represent State primacy agency SDWA monitoring results as provided by the States. The data represent the analytical results for Community Water Systems (CWSs) and Non-Transient Non-Community Water Systems (NTNCWSs) that are required to monitor for the Phase II/V chemicals. Some States included data from Transient Non-Community Water Systems (TNCWSs), but these systems are not required to monitor for the Phase II/V contaminants. The TNCWSs' data are almost exclusively for nitrate, which is not included in these discussions. Some State data sets are not complete for all systems. Massachusetts had only VOC data available from approximately one-third of their CWSs. The Ohio database only includes results from a special study of selected herbicides (SOCs) in surface water systems (Ohio makes these data available on an internet site, http://www.epa.state.oh.us/ddagw/pestspst.html). Michigan provided an excellent database for SOCs and VOCs (volatile organic chemicals), but it does not include IOC data (inorganic chemicals).

States provided these data sets voluntarily; some were available publicly; and a few were requested to broaden the spatial coverage of the analysis. For example, Iowa and California have published reports of their analysis of SDWA data (Hallberg et al., 1996; California DHS, 1997). California also makes their data publicly available on CD ROM. But for purposes of this analysis (so data could be analyzed in a consistent manner among states for compatibility) it was necessary to get a copy of the data in a different format.

Table II. Principal State and supplemental databases used for analysis in this report.

State	**Contaminant Groups Represented (IOCs, SOCs, VOCs, O=other contaminants)**[1]	**Number of Analytical Results included in this Analysis**	**Number of PWSs with Results for this Analysis**	**Time Period**
Alabama	IOC, SOC, VOC, O	708,569	731	1985-1998
California	IOC, SOC, VOC, O	3,897,362	6,414	1984-1998
Illinois	IOC, SOC, VOC, O	2,967,946	1,392	1987-1997
Indiana	IOC, SOC, VOC, O	257,428	1,488	1982-1997
Iowa	IOC, SOC, VOC, O (published report)	458,320	2,090	1988-1995
Massachusetts	VOCs only	76,737	322	1993-1997
Michigan (CWS)	SOC, VOC, O	346,181	1,146	1993-1997
Michigan (NTNC)	SOC, VOC, O	339,540	2,106	1993-1997
Montana	IOC, SOC, VOC, O	276,675	1,786	1993-1998
New Jersey	IOC, SOC, VOC, O	980,915	4,503	1993-1998
New Mexico	IOC, SOC, VOC, O	266,262	1,299	1992-1996
Ohio	SOC (only selected herbicides in SW systems)	6,646	144	1996-1997
Oregon	IOC, SOC, VOC, O	169,521	2,345	1990-1998
STATE TOTALS		**10,752,102**	**25,766**	
URCIS (40 States & Territories)	some SOC, VOC, O	3,492,480	24,357	1983-1992
Novartis (21 States)	SOC (atrazine and simazine)	128,603	15,766	1993-1996
AWWSC (19 States)	IOC, SOC, VOC	55,526	137	1995-1996
USGS (42 States)	some SOC, VOC, O			1992-1997
TOTALS		**14,428,711**		

---

[1] Includes data for: IOCs = the 13 regulated inorganic chemicals in Table I.A.1; SOCs = the 30 regulated synthetic organic chemicals in Table I.A.1; VOCs = the 21 regulated volatile organic chemicals in Table I.A.1; Other = other regulated or unregulated chemicals not listed in Table I.A.1.

Table II also summarizes four other major sources of data. Coupled with the State information, the data reviewed for this report include more than 14 million analytical results. (The number of systems from the States and these supplemental sources is not totaled because there is overlap among these special databases and some States, and hence many systems would be double-counted.) The supplemental databases include:

- EPA's Unregulated Contaminant Monitoring Information System (URCIS). Includes data from about 40 States and Territories from the first round of unregulated contaminant monitoring (1989-1993). These data also include samples taken prior to 1989 which were grandfathered into the database. (The values shown here for URCIS may differ from other EPA reviews because of the screening conducted for this study to ensure consistency in analysis.)

- A special regional analysis of atrazine and simazine occurrence in CWSs, conducted by Novartis Crop Protection, Inc., the pesticides' registrant. This study was conducted as part of a special study under direction from EPA's Office of Pesticide Programs. While limited to two contaminants, this study is important because it compiled data from CWS monitoring conducted between 1993 and 1996, from 21 States across the country, for one of the most widely occurring contaminants. It represents one of the most complete assessments of drinking water occurrence data of its kind (Clarkson et al., 1997).

- Drinking water data from systems operated by the American Water Works Service Company. AWWSC operates systems in many states, and while it keeps relatively few records (compared to State occurrence databases), its multi-state nature helped to provide insights early in this analysis, and to serve as a comparison to the representative cross-section results. (AWWSC also collects various supplemental data and maintains its own data system, which provided useful insight on various data quality issues.)

- Studies of the U.S. Geological Survey (USGS), particularly results from the ongoing National Water Quality Assessment Program (NAWQA) and other special contaminant studies (e.g., the Mid-Continent Herbicide Study, the Toxics Program). The USGS conducts extensive monitoring activities under nationally consistent protocols. This makes the data very valuable for reviewing national patterns of occurrence that can provide perspectives on drinking water problems. As further discussed in this report, the majority of these data are from ambient water sources and *not* drinking water. In all, more than 75 USGS reports on surface and ground water quality in parts of 42 States were reviewed. (A listing of USGS reports reviewed is given in Appendix B.)

II.A. Data Management

There are numerous data handling and management issues, as well as data quality issues, that had to be addressed to enable the analysis presented in this report. The primary objective was to develop a consistent and repeatable approach that would allow valid comparisons between and among the various data sets, allowing the data to be jointly evaluated to provide an overview of occurrence patterns at the national level. Some of these issues are reviewed below as a preface to understanding the results.

In general, States and other sources did not reorganize or reformat data, but simply transmitted the data in whatever manner was easiest. For example, while the Phase II/V compliance data from 1993-1995 were of greatest interest, in many cases it was easier for the State simply to transmit their entire data set, which generally contained information on all chemical contaminants, over a greater span of years (as noted Table II). In addition to the regulated contaminants, every database included data on additional contaminants which varied among the databases.

Data were transferred using three main media: FTP, e-mail, and diskettes (including zip-disks or CD-ROM). Evident from the number of analytical results presented on Table II, these are very large databases, often several megabytes in size, and transmission was often complicated. Many of the data sets were received "as is" and had not been formatted by the State in any way. The data were received in a number of file types including spreadsheet files, DBF files, and THM files. Each database was unique in format, layout, custom codes, and data element usage.

After receipt, an initial review of the information in each database was performed. In most cases, the data was not accompanied by a protocol outlining each variable. In many cases, the variable headings could be determined by examination. In every instance, follow-up with the State/data source was necessary to decipher variable headings or contaminant codes. When all variables were understood, a formatting plan for the data was developed. Nearly all of the data sets required some type of formatting to facilitate analysis. Data formatting problems varied from one dataset to another. (Some of the most common data problems are outlined in Appendix C.)

All statistical analyses were conducted in SAS® statistical software. Data formatting problems were corrected in Microsoft® Excel with the aid of specialized programs written in Visual Basic® or were corrected directly in SAS before the analysis began.[2] Data formatting was the most time consuming and labor intensive part of the data analysis. Each data set presented unique challenges. While analysis of the data was consistent from one data set to another, each data set required some unique editing and filtering because of differences among basic data elements (as described below).

II.B. Data Elements

Each drinking water contaminant database was reviewed to ensure it contained the basic data elements (data fields) necessary to conduct a consistent analysis for this study. These elements were reviewed with State/data-source contacts both before and after data were received to ensure consistent and appropriate interpretations. While the presence of such elements enables the various data sets to be analyzed in a similar manner, each also may be used in unique ways by the individual States/sources. A brief review of these elements, examples of some of the common problems encountered, and some of the data editing that was required, are also presented.

---

[2] SAS is a registered trademark of the SAS Institute, Inc. Excel and Visual Basic are trademarks of the Microsoft Corporation.

II.B.1. Analytical Results Data Elements

Contaminant Identifier (ID). States use different identifiers including EPA OGWDW Contaminant IDs, STORET contaminant ID numbers, and unique State contaminant identifiers, and/or some combination of the above. Some include numeric identification codes as well as the chemical or trade name of the contaminant. In some cases, multiple IDs referred to the same contaminant. In most cases it was necessary to confirm the "translation" of IDs to particular contaminants. CAS numbers were added to ensure accurate identification and processing. Crosswalk tables had to be constructed to ensure that all entries could be converted to uniform IDs.

Analytical Method. In general, this was entered as the EPA method number used for the analysis, but this also varies among States, with some using their own unique method IDs. In several cases there was no record for this element. At the State level, the assumption was that the State certified labs use only approved methods, so this field was not necessary for every sample.

Detection or Reporting Limit. Knowing the Minimum Reporting Level (or Limit, MRL) for a contaminant-method combination can be very helpful for understanding some aspects of the results. Very few databases contained a field for this element. A few provided either a crosswalk table or a general paper copy. (The MRL can sometimes be surmised from the data, if less-than fields and data values are recorded. Also, the MRL is commonly referred to as the Method Detection Limit (MDL) but this is not strictly correct. The MRL is the minimum concentration that can be reported with confidence and is based on the MDL. The MDL is a function of the method and equipment used to determine the presence of an analyte in relation to background noise. The MRL is often 5 to 10 times greater, or a statistical multiplier greater, than the MDL.) Although most databases did not provide the MRLs, some basic information is available for compliance analyses. Within for the Drinking Water program EPA specifies MDLs (MRLs) for all VOCs; for SOCs trigger levels are specified that are essentially multi-method practical quantitation limits that serve as MRLs for most States. IOCs are the most variable because of the diversity of methods in use.

Analytical Result's Elements. The actual analytical results generally comprise multiple fields: sometimes the sign of the results; sometimes the unit of measure; and the actual numerical value of the analysis. One critical component of the results sign in such databases is how "<MRL" (less-than) values are recorded. This affects fundamental data processing procedures before any statistical processing can even begin. State databases vary widely in how values less than the reporting level were recorded. Some record a "<" symbol in a discrete field; others simply enter zero in the results field; some entered ND or another code; many had mixed usage that had to be carefully resolved. In most cases, resolving these issues required careful, time-consuming editing. (How less-than values are handled is a significant issue for any parametric statistical calculations. In this report we focus on simple non-parametric approaches which are not hampered by these problems.)

Other signs are also used (+, -, =, and other codes) related to unusual analytes or unique codes that identify an aspect of the data. In a few cases, unique codes were identified that the State could not confirm; such data were eliminated.

The other fields were carefully reviewed to ensure that all results were entered in (or were converted to) the same units of measure. For this analysis, all results were processed and are reported in mg/L (parts per million). State databases handled this in different ways, also. Some recorded a results field using both mg/L and µg/L, or other units as appropriate; some States recorded a uniform string of numerals with a code for where the decimal place should be located; while some entered all data in the same units and had no record; and other States assumed that certain contaminants were always entered in the same units (e.g., IOCs in mg/L, SOCs in µg/L). Again, careful review was required to ensure correct units were used. In nearly all databases, some small number of values appeared that were significant outliers, suggesting that a unit conversion or original reporting problem was likely. In some cases, these could be confirmed with the State and then eliminated. These data were rare and do not affect any general statistics unless the maximum concentration value was of interest. In one case, for a few analytes, a State knew aberrant data existed (because of a reporting error), and results greater than the 99th percentile concentration were deleted.

II.B.2. System and Sample Elements

Public Water System Identifier (PWSID). Most States did use the SDWIS PWSID but some used unique State identifiers, while others used both. In some cases, crosswalk tables had to be constructed from other State databases to ensure consistency. Many databases also included the system name, which was helpful for cross-checking problems with system IDs (e.g., when an ID was entered incompletely or in error). In a few databases, elements were included identifying a particular facility or treatment plant belonging to the PWS.

Source Water Type. The source of the raw water was identified in all datasets. The identification was typically by a code for surface water, ground water, both, purchased, or other (and a myriad of minor codes, unique to different States). For this analysis a primary objective was to subdivide and summarize data by systems using surface water and ground water. Systems using both, and systems using ground water under the influence of surface water, were included with surface water data. Systems classified as purchased or consecutive, and those with no source identified, or “other” sources were not included in the analysis. This was a small component of the data.

Other System Features. A few States include other information in their databases that were used in the analysis (or could be useful in future analyses). These included data on the population served by the system, system type (CWS, NTNCWS), or other locational data such as latitude and longitude.

Sample Identifiers. Sample IDs and other sample elements were also handled in a variety of ways by the States. Many desired elements were included in some databases, including: a unique sample ID (number); sampling point identifier (name or number for a specific site where a sample was collected); sample location type (e.g., entry point to distribution system; distribution line; well head); sample type (e.g., compliance, confirmation, duplicate, special, finished/treated, raw, and many others); sample collection date; and sometimes date of analysis. Sample location type and sample type were often a single field, with limited information. On the minimal side, sometimes only the sample type and date of sampling were included. Date had to be combined with PWSID to form a discrete sample ID to facilitate these analyses. The dates were also used in conducting various temporal analyses and in correlating samples to assess co-occurrence of contaminants.

II.C. Data Quality and Consistency Issues

There are a myriad of data quality issues involved in a project such as this. Many have been noted in the preceding discussions. A few key issues are reviewed here in further detail.

As noted, every State database contained unique data elements or unique treatment of common elements. Even after initial screening and conversion, unique factors were always uncovered during data analysis. These were resolved in consultation with the State/source. As a general rule, when errors or ambiguities in various data elements could not be resolved, the result was eliminated from the analysis to avoid aberrant results. This eliminated very few observations (compared to the thousands of analytical results included in the databases).

There are always concerns with the quality of the analytical data entered in such databases. This study only used data from State primacy agencies -- official data from the regulated drinking water program. Laboratories are certified for drinking water programs, which assumes the use of various quality-control and quality-assurance procedures. Certainly problems exist, but the base assumption is that the analytical data are of good quality.

When computing basic occurrence statistics, such as the number or percent of samples or systems with detections of a given contaminant, the value of the minimum reporting level (concentration) can have important consequences. The lower the reporting limit, the greater the number of detections. Within reason, MRLs can even vary from lab to lab using the same method, or can vary with sample batch, etc. There can be more dramatic variation if different methods are used for the same contaminant. Within the drinking water program, methods have become well standardized so this was not a significant issue for this study, particularly for the SOCs and VOCs. For example, the VOCs EPA requires an MRL (MDL) of 0.5 μg/L. For SOCs, trigger levels are specified that functionally have ben used as multi-method MRLs by many States. The compounds most affected are the IOCs, where many more methods are available, with a much greater variance in MRLs. (This will be discussed further in the review of results.) MRLs are not specified in many State databases, as discussed. Hence it was not possible to “standardize” the detections to a fixed MRL.

Another issue that can affect a large-scale summary of results, is the different sampling schedules that may be used by different PWSs. A PWS with a known contaminant problem usually has to sample more frequently than a PWS that has never detected the contaminant. Obviously, a simple computation of the percentage of samples with detections (or other statistics) can be skewed by the more frequent sampling at the contaminated site. This analysis is focused on occurrence at the system level, which avoids the skewness inherent in the sample data, particularly over the multi-year period covered in the summary statistics. Also, as noted above, the data used in this analysis were standardized to use only data with confirmed water source and sampling type information. Only standard SDWA compliance samples were used; “special” samples, or “investigation” samples (investigating a contaminant problem that would bias results), or samples of unknown type were not used in the analysis. Various quality control and review checks were made of the results, including review by the States providing the data. Many of the most problematic data quality problems encountered occurred with older data (especially, pre-1990-1993). In some cases, as noted, these were simply eliminated from the analysis.

Another data management issue involves short-hand codes used for the most fundamental data element: the analytical results. Many States have special codes they enter for multiple contaminants that are covered under the same analytical method. One of the most common approaches summarizes the results of a single method with an 'ND' or zero (for all contaminants not detected) and individual observations only for those contaminants with a positive result. This system is best illustrated using the 21 VOCs covered under method 502.2 as an example. If none of the 21 VOCs were detected, a State might enter '21 VOCs' in the contaminant ID column and '0' in the results column, as if it were a single observation. If one or more of the 21 VOCs were detected, these individual contaminants would be entered in the contaminant ID column and the concentration detected would be entered in the analytical results field as individual observations. It then is assumed, in this case, that the other contaminants were tested for but not detected. While this system works well for data storage, it does not work well for data analysis. To statistically analyze the occurrence of contaminants a result is needed for every observation (sample). Hence, special processing was required for every data set to unravel these codes. When "21 VOCs" were 0, the contaminant ID and the 0 actually had to be entered in the dataset for all 21 of the contaminants, to fill out the record. Likewise, when two of the contaminants from the VOCs were entered, the processing had to recognize which two, and then fill out the dataset for the additional 19. Although the example above may seem straightforward, the reality of implementation was often more complicated. It was not unusual to find caveats in the database which needed to be further defined by the State before analysis could begin (e.g., sometimes the rules for processing such entries changed over time). This is done differently among States, and it is typically is not entirely consistent within a State dataset.

All of these controls and data editing were performed to make the results as reliable and consistent as possible and to ensure that they were unambiguously from finished drinking water, representing the product delivered to the public. Also, elimination of data with inconsistent elements helps to ensure that the analysis is relatively repeatable, for future consideration.

II.D. URCIS Data Quality

As noted, for this study we also reviewed some results from the URCIS database. There are data quality issues unique to URCIS. These are reviewed to put URCIS results, and their limitations, in perspective. EPA's Unregulated Contaminant Monitoring Information System (URCIS) is a compilation of public water system monitoring results for contaminants without NPDWRs (i.e., unregulated contaminants) that are also collected under the authority of SDWA (some of the contaminants included in early rounds of this monitoring later became regulated contaminants with NPDWRs). EPA collected these data from States and has been working on the data clean-up and analysis since 1992. The version of the database used here was a current edition of data (from 1997) selected for other on-going analyses. The values shown here for URCIS may be somewhat different than for other EPA reviews because of the screening conducted for this study to ensure consistency in analysis. The data includes VOC and some SOC results from 39 States and the Virgin Islands.

Data from URCIS used in this analysis range in date from 1983 to 1992. The majority of the data are from the first round of required, unregulated contaminant monitoring that began in 1987. Because of the age of these data (in relationship to the rapid improvements that have been made in data processing systems) the quality of data received by EPA for URCIS was highly variable. The data were supplied by States on a variety of media, ranging from photocopies of hand-written files to electronic files on magnetic tape or diskette, and in many different formats and software configurations. Some data were electronically transferrable, other data had to be manually entered or re-entered; legibility was occasionally a problem in entering the data, introducing the possibility of transcription errors.
The sampling point designation in the original data is often unclear or omitted and many samples are not identified as coming from the source water, an entry point, or elsewhere in the distribution system. It is not always clear whether a sample was of raw or finished water. The method of analyzing samples was not always reported or was reported incompletely or in a code other than EPA method numbers. Not all samples had certification number for laboratories conducting the analyses. In some cases, the system water source (surface or ground water) was not included, but most of these elements have been completed by matching URCIS records with SDWIS.

All of the original data entered into the URCIS database had the concentration units converted to μg/L; some data sets, however, included data with mixed units (μg/L and mg/L) and the conversion resulted in abnormally high concentrations for some observations. While the observations with the highest concentrations were excluded from the analysis as suspect, there may be observations remaining in the database with values up to three orders of magnitude greater than the actual concentration. While these observations affect a review of the maximum concentration values of a contaminant, and other parametric calculations (e.g., means) they will have limited impact on the important occurrence statistics reviewed in this report. As with the State SDWA data, entries without source water codes or other key elements, or those having undecipherable codes, were eliminated from the analysis. There are other features of the data that also must be understood.

Results from seven States in URCIS include only chemical analyses with positive detections. Results from six other States are all between 88% and 99% detections; results for these States are probably all from multi-analyte samples with a positive detection for at least one analyte. As a result, detection rates in URCIS are biased toward a greater percentage of contaminant detections than is actually the case. In addition, the full range of contaminants tested for is not discernable for those States which reported mostly detections instead of all results. However, for States that submitted complete data, the URCIS results are very consistent with the results from the complete State SDWA databases collected for this study (for the analytes they have in common).

In spite of these shortcomings, URCIS is such a large database, with data from so many States, that it has considerable value. The sheer quantity of data in URCIS somewhat compensates for some of the quality problems, and the spectrum of coverage (40 States and Territories) warrant review. Further, most of the data quality problems bias the data in the same, conservative direction: the occurrence of contaminants is likely over-stated in both frequency and maximum concentrations. Thus, aggregate values should not underestimate occurrence.

II.E. Basic Analysis and Review

As noted, the basic data manipulations and statistical analyses were all performed using SAS programs. The basic data summaries output for each of the datasets included the following statistics for each analyte (and some groups of analytes) by ground water and surface water:

- Number of samples
- Number and percentage of samples with reported values greater than the MRL
- Number and percentage of samples with a reported value greater than half the MCL
- Number and percentage of samples with a reported value greater than the MCL
- Number of systems (sampling for a given analyte)
- Number and percentage of systems with reported values greater than the MRL
- Number and percentage of systems with a reported value greater than half the MCL
- Number and percentage of systems with a reported value greater than the MCL
- The concentration of the median, 95th, and 99th percentile (in mg/L)

Note that a reported anayltical value greater than the MCL does not constitute an MCL violation. An MCL violation occurs when the MCL is exceeded by the average of four quarterly samples or confirmation samples as required by the primacy State. Occurrence values were calculated based on the number of samples, primarily for evaluating the data. The results described in this report are those calculated on a system basis (i.e. the percentage of systems with detections) to avoid the bias and skew inherent in the sample data.

The results were reviewed and cross-checked with some duplicate programming. Results were compared with published or unpublished analyses from the states, when possible. Summary reports and data tables were returned to the States/sources for their review and comment.

III. TOWARD A REPRESENTATIVE SAMPLE

As discussed, this evaluation is one of the first and most extensive analyses EPA has conducted on the occurrence of chemical contaminants regulated under the Safe Drinking Water Act. One primary purpose is to begin to provide better data to address issues related to occurrence and monitoring of public drinking-water systems. Another objective, even at this initial stage of analysis, is to provide an overview of occurrence patterns for the 64 regulated chemicals at the national level. Currently, there is no complete record of analytical results collected under SDWA that can be processed for a comprehensive national overview. EPA's Safe Drinking Water Information System (SDWIS) maintains a variety of water system inventory and operation information as well as compliance program information. For most contaminants, the only analytical results filed in SDWIS are those related to violations of an MCL. The analytical results from monitoring of the Phase rule chemicals, and most other contaminants, are stored in the individual State databases. There has been no feasible way to access these data to construct a nationally representative sample except, as undertaken here, through analyzing the individual State datasets.

As noted, most of the data for this review were provided voluntarily by States. Obviously, constructing a "representative" view from such data can be problematic, so additional data were requested from a few States to broaden the spatial coverage of the analysis. The 14 States whose data were reviewed are shown on Table III (the 12 States whose data were analyzed in detail are shown on Table II). While 14 of 50 States is a substantial sample, it is not necessarily representative. The data from all these States could simply be aggregated to compute a composite occurrence value for a contaminant. However, the result would significantly over-represent Midwestern "Cornbelt" States (i.e., the sample could contain Kansas, Missouri, Iowa, Illinois, Indiana, and Ohio, over 40% of the States represented). Hence, various means were evaluated to enable the construction of a cross-section from the available State databases that would provide a reasonable first view of national occurrence.

There are many sophisticated statistical methods that can be applied to analyze limited (and biased) data. However, at this early stage of evaluating SDWA occurrence data this did not seem warranted. For initial analysis, we attempted to develop a simple approach, that would be clear and repeatable, resulting in aggregate numbers that could be easily understood. When data from more States can be compiled, more involved methods may be warranted, especially after some initial determination of the variance for the different contaminants.

III.A. Representativeness

Two broad factors were considered in the assessment: geographic or spatial diversity — the wide range of climatic and hydrologic conditions across the United States must be represented to some degree; and pollution potential — the States should represent the range of the likelihood of contaminant occurrence. Many past EPA studies have shown that some simple measures, such as population (or population density) are valid indicators of pollution, because it is human activity and land use — be it manufacturing or agriculture — that is the source of most pollutants, particularly the organic chemicals. Various demographic or other factors were evaluated as independent measures or indicators of pollution potential. In general, manufacturing/industrial activity, as well as population density, are considered the major sources of many VOCs (degreasers, solvents, petroleum compounds). Most SOCs are pesticides, and agriculture is the largest user of these compounds. While IOCs have various uses in manufacturing, they also occur naturally and ambient concentrations can be enhanced by mining or other diffuse activities. Natural geologic sources of IOCs were not directly considered in the assessment for representativeness, in part because whole States needed to be evaluated and such sources are often localized. However, by including geographic or spatial coverage across the Untied States as a factor (e.g., from Illinois to Montana) a range of geologic conditions are inherently included (as well as the hydrogeologic and climatic variability).

More than thirty-five different factors that might be useful indicators of pollution in each State[3] were considered in this study. These ranged from Census data on manufacturing, agriculture, and population density, to indices such as EPA's Section 106 allocation factors or the *1991-1992 Green Index: A State by State Guide to the Nation's Environmental Health* (prepared by the Institute for Southern Studies). Two methods were considered for evaluating the States' comparable pollution potential. The first was development of a singular numerical index, incorporating factors such as manufacturing in the State, total pounds of chemicals released, and pesticides used, into a comprehensive ranking for each State. However, such a ranking for all sources requires various factors to be weighted, and the meaning of the resultant number can be difficult to understand, as well as argumentative.

For this initial analysis, a second, simpler method was used to evaluate the pollution potential and the representativeness of the States. One primary factor was chosen to indicate the potential pollution from manufacturing and one factor for agriculture in each State. States were then ranked from 1-50 for each factor and divided into quartiles based on the ranking. The rankings were reviewed to assess if States could be selected in approximate balance from each quartile. In addition, a number of secondary factors were also considered to further insure that the data were representative. This method does not, of course, avoid all of the problems discussed above, but it does provide a simple look at the conditions represented by the States. The resultant rankings are summarized on Table III; the factors considered are discussed briefly, below.

III.A.1. Manufacturing Indicators

Numerous factors were considered as potential indicators of manufacturing-related pollution, including EPA's Toxic Release Inventory (TRI) (including total releases, releases per square mile, and releases excluding air releases), the number of manufacturing establishments, the number of manufacturing establishments per square mile, the number of manufacturing employees, the value added by manufacturers, and the value added per capita (*Annual Survey of Manufacturers,* 1995; *Census of Manufacturers,* 1992; and *Toxic Release Inventory,* 1995). These factors were each considered in terms of their inherent value as pollution potential indicators, their range and variance (in providing a relative ranking of the States), and their inter-relationships.

The total TRI releases per square mile, number of manufacturing establishments per square mile, and value added per capita were considered the three most useful indicators. The TRI was considered useful because it is a measure of how many pounds of toxic chemicals are released within the State. While there are problems with the TRI (e.g., some inconsistent release estimation techniques; omission of many small establishments, or those with releases below specified thresholds), it can validly be used as a direct indicator of potential pollutants released. The number of manufacturing establishments per square mile takes into account how many factories are actually engaged in manufacturing and thus how many establishments potentially

---

[3] The data are analyzed on a statewide basis, thus any determination of representativeness must be based on whether the States for which information is available are representative of the nation as a whole. There are, of course, problems with using States to determine representativeness since States are large, diverse entities; however it is not practical to break the data down any further.

contribute to pollution. By breaking down the number of manufacturing establishments per square mile, the size of the State is also taken into account. The final factor that was considered to be viable was the value added by manufacturers per capita. This is an attractive measure because the value added should correlate well with how much is actually produced within the State. The problem with the measure, and this is also a problem with the number of manufacturing establishments per square mile, is that it does not take into account the variation in pollution released by different industries. For example, an industry that adds a lot of value to a product may cause little pollution while another industry that does not add much value may contribute more pollution.

The data clearly show a close correlation between the number of manufacturing establishments per square mile and the population density in each State, as well as a clear linear association with the total TRI pounds released/square mile, number of manufacturing employees, and total value added. Hence, the number of manufacturing establishments per square mile was used as the primary indicator. The other key reason for choosing this factor was that it is a simple measure of how many establishments are actually engaged in manufacturing and thus are potentially polluting sources of drinking water. The TRI total pounds released per square mile was used as a secondary factor in determining representativeness. Squillace and others (1999), from the USGS, found a significant correlation between VOC occurrence in ambient ground water and population density, in a national NAWQA study. As noted, population density and manufacturing density are highly correlated. Manufacturing density and TRI data were used in this ranking because they were considered more direct measures of pollution potential for this study.

III.A.2. Agricultural Indicators

There is no complete measure of pesticide usage by States that is readily available. So, a variety of factors were considered to assess potential organic chemical pollution from agriculture in each State. These included the percent of the State's population that is classified as rural, the percent of land in the State that is cropland, the percent of land that is grassland pasture and rangeland (a possible inverse indicator), and total farm agricultural chemical expenses. Like the manufacturing factors, these agricultural variables were considered in terms of their value in indicating potential sources of pollution and were plotted against one another to determine how closely they are correlated.

Of these factors, total farm agricultural chemical expenses was considered to be the best indicator of potential pollution. The percent of the State's population that lives in rural areas does not necessarily relate to agricultural chemical use or cropland. There is, of course, a correlation between cropland and agricultural chemical use, but there are notable exceptions such as Florida and California which use a large amount of agricultural chemicals despite having more limited cropland area. While there are some incomplete surveys of pesticide use, the *Census of Agriculture* (1992) measure of dollars spent on agricultural chemicals, is a more consistent and complete measure.

III.B. Representative Cross-Section of States

Table III summarizes the quartile rankings of the States selected to approximate the national cross-section used in this summary report. Eight States were selected, from the 14 whose data were evaluated, as providing the best balanced cross-section, based on geographic coverage, relative rankings for pollution potential (i.e., potential for contaminant occurrence), and data quality and completeness. These States provide representation from across the U.S., from Alabama to Oregon to New Jersey. They provide some representation from all quartiles of pollution potential indicators, and, hopefully, a balance in potential occurrence. The data from these eight States were used to compute aggregate occurrence values (i.e., the percentage of water systems that had a detection of contaminant X) as an approximation of a national cross-section. For balance, all the SDWA data (see Table II) were used to show the range of occurrence in the results summaries. However, the ranges for only a few of the 64 contaminants were extended by the additional data. While the data from these Cross-Section States cannot be stated to be “statistically representative,” their distribution should provide a clear indication of central tendency. The results will be compared with other data for perspective and as a comparative check. It is also worthy of note that the 8 Cross-Section States represent over 25% of the U.S. population using PWSs, and over 20% of the PWSs, a substantial sample. (Considering the other States evaluated and used in the range estimates, the data include over 40% of the population and nearly one-third of PWSs.) Again, if additional State databases can be added to this analysis, other statistical techniques can be utilized to develop national estimates of occurrence.

Table III. States with water quality data included in the drinking water occurrence assessment. States are listed by their national rank order (1=highest, 50=lowest) for the number of manufacturing establishments per square mile. The rank for TRI releases in pounds per square mile is in parentheses. Also listed is the State's rank for total expenditures for agricultural chemicals. States on the left are those selected for the national cross-section compilation; States on the right are others whose data were evaluated and are included in the data ranges summarized.

Quartiles for rank-order of all States	National Cross-Section States	National Rank: Number of Manufacturing Facilities per square mile	National Rank: Total Expenditures, Agricultural Chemicals	Other States; Data Used for Ranges and Check on Cross-Section	National Rank: Number of Manufacturing Facilities per square mile	National Rank: Total Expenditures, Agricultural Chemicals
1	New Jersey	2 (8)	37			
	Illinois	10 (11)	2	Massachusetts	4 (22)	43
	California	11 (38)	1	Ohio	6 (2)	11
2	Michigan	13 (16)	18	Indiana	15 (6)	7
	Alabama	25 (7)	26			
3				Missouri	26 (26)	12
	Oregon	34 (39)	22	Iowa	36 (28)	3
4	New Mexico	44 (40)	40	Kansas	39 (37)	16
	Montana	48 (34)	34			

IV. THE SDWA OCCURRENCE DATA

In the following sections of this report we summarize the results of the analysis of the drinking water contaminant data. The summary discussions will focus on basic occurrence patterns, and discuss the percentage of water supply systems in the national cross-section (or individual states, to illustrate particular points) that had detections of a given contaminant greater than the MRL (>MRL; simply put, systems with detections), systems with detections greater than 50% of the MCL (>0.5MCL), and systems with detections greater than the value of the MCL (>MCL).

In this study, data were analyzed for approximately 280 different contaminants. This includes the 64 contaminants regulated under the Phase rules (Table I.A.1), six other regulated contaminants (e.g., nitrate, lead and copper), 34 unregulated contaminants that were formally required for monitoring under past regulations, 14 contaminants that were discretionary for monitoring under these regulations, and approximately 160 other unregulated contaminants. Of these other unregulated contaminants, 22 are on EPA's Contaminant Candidate List, including 16 that appear on the proposed new unregulated contaminant monitoring list. The remaining 140 are other unregulated contaminants that were contained in the State databases. Beyond the contaminants that were required for monitoring under SDWA, the data are very incomplete, sometimes consisting of only a single analysis. Of the 10.7 million analytical records processed, about 73% were for required monitoring.

As noted earlier, data were processed from approximately 26,000 water systems. This does not mean that there are data from 26,000 systems for each contaminant. As discussed, not every State provided data for all systems or all contaminants. Even within a State, some systems may have waivers for particular contaminants or groups of contaminants. The maximum number of systems for a particular contaminant is approximately 16,000 for many VOCs and SOCs. For the regulated contaminants the number of analyses varies from about 50,000 to 150,000 for each contaminant.

The narrative summary will not attempt to discuss results for all 280 contaminants or even the 64 Phase rule chemicals. The summary will focus on illustrating key issues and answering key questions. Most of the summary tables will focus on a few representative IOC, SOC, and VOC contaminants. (More complete summary tables are presented in Appendix A.) In addition to data for individual contaminants, summary data will be presented by contaminant groups.

IV.A. Contaminant Groups

The group classifications are partly related to the nature and source of the contaminants and partly to their chemical properties and general method of analysis. The aggregate groups provide additional perspective because many of the compounds are inter-related by source or fate and transport characteristics. For this analysis, the occurrence of any contaminant in the group is counted system by system. A system may detect one or all contaminants in the group, but is only counted once. The analysis by contaminant groups provides an indication of the aggregate occurrence and some insight into co-occurrence. The contaminants are grouped by their standard categories, IOCs, SOCs, and VOCs, and also into smaller categories.

IV.A.1. SOC Groups

The SOCs are divided into high-occurrence SOCs (HiSOCs) and low-occurrence SOCs (LoSOCs). The HiSOCs and LoSOCs Groups were used for the impact analysis estimates of the draft CMR. SOCs Groups 1, 2, and 3 are further subdivisions that will be used for the current analysis. SOCs Group 1 aggregates the data for the commonly occurring herbicides that often dominate the occurrence of HiSOCs; these are alachlor, atrazine, and simazine. These three contaminants are often analyzed from the same sample (using a multi-analyte method). SOCs Group 2 includes adipate, phthalate, and benzo(a)pyrene, which vary in occurrence. In some studies, phthalate and adipate may have high rates of occurrence, but this has, in part, been attributed to sample contamination resulting from the use of plastic piping or even plastic sampling and lab equipment. SOCs Group 3 includes the remainder of the SOCs (with the exception of the PCBs and Dioxin). These tend to be either low or localized in occurrence, related to the contaminant’s chemical properties and regional use patterns.

IV.A.2. VOC Groups

While all of the 21 currently regulated volatile organic compounds (VOCs) are similar (i.e. they evaporate readily), it can be useful to differentiate between two smaller groups of VOCs: VOCs Group 1 and Group 2. While not all of the 21 regulated compounds are included in these subclassifications, distinguishing between these two groups is very helpful, as many of the compounds within each group have similar environmental sources and fates.

The Group 1 VOCs include the lighter, non-halogenated compounds: benzene, toluene, ethylbenzene, and the xylenes. Aside from having many structural similarities (all contain a benzene ring substructure), all of these compounds are either natural components or derivatives of crude oil (Howard, 1997). Thus, major environmental sources of these compounds are gasoline and other petroleum fuels. Other potential sources for Group 1 VOCs are industrial solvents, carriers or solvents in pesticides and paints (Howard, 1997). Because of their low water solubilities (ranging from 0.16 to 1.78 g/L at 25°C; Bloemen and Burns, 1993) and low densities (ranging from 0.86 to 0.88 g/cm^3 at 20°C; Lide, 1996), these compounds often form light non-aqueous phase liquids (LNAPLs) which tend to "float" on the water surface.

As gasoline and other petroleum fuels are common sources of Group 1 VOCs related to fuel spills and surface run-off, these compounds are more likely to be found in surface waters than in ground waters. Because Group 1 VOCs share a potential source, the likelihood of Group 1 compounds co-occurring with other Group 1 compounds is greater than with other regulated VOCs, as none of the other 17 regulated VOCs are commonly found in gasoline (Howard, 1997). In surface waters, there is often a pattern of frequent Group 1 VOC occurrence, but at low concentrations. Also, as "floaters", when they occur in ground water they tend to be found more often in shallow ground water than in deeper ground water (Hallberg et al., 1996). It is possible that this pattern is at least partially attributable to the ubiquity of their environmental sources coupled with the fact that they form LNAPLS, which facilitates their volatilization, and leads to lower concentrations in surface waters. Also, these lighter compounds are more subject to biological degradation and there is greater residence time in the soil-ground-water system allowing for biological processing.

The chemicals which compromise the Group 2 VOCs are the heavier, halogenated organic compounds. The six compounds included in this group are tetrachloroethylene (PCE), trichloroethylene (TCE), *cis* and *trans* 1,2,-dichloroethylene, 1,1,-dichloroethylene, and vinyl chloride. As with Group 1 compounds, all of these compounds are structurally similar: they all have a double-bonded, two-carbon-chain backbone, and all contain at least one chlorine atom. None of the Group 2 VOCs are known to occur naturally. While many of these compounds are used as metal degreasers and as industrial solvents, some may also be used for dry cleaning (PCE: Howard, 1997; TCE: Parsons, et al., 1984), in paint and ink formation (TCE; Howard, 1997), and in synthetic polymers such as PVC (vinyl chloride; Howard, 1997) and plastic wrap (1,1-dichloroethylene; Howard, 1997). As with Group 1, Group 2 VOCs tend to form non-aqueous phase liquids because of their low water solubilities. However, with the exception of vinyl chloride, Group 2 compounds tend to "sink" in the water column because of their higher densities (ranging from 1.21 to 1.62 g/cm^3 at 20°C; Lide, 1996), and thus are classified as *dense* non-aqueous phase liquids (DNAPLs). When present in ground water, this behavior provides limited access to the air/water interface, and thus these compounds may tend to volatilize less than Group 1 VOCs.

Although each Group 2 compound may have a discrete environmental source, many of the compounds are components of the PCE degradation series. Upon its release into the environment, PCE may undergo microbial degradation, losing chlorine atoms one by one and forming first trichloroethylene (Parsons et al., 1984), then the dichloroethylenes (*cis*- and *trans*-1,2-dichloroethylene: Parsons et al., 1984; 1,1-dichloroethylene: Howard, 1997), and eventually vinyl chloride (Howard, 1997). Additional degradation paths are possible, some eventually leading to carbon tetrachloride and dichloromethane, two other regulated VOCs.

Because of their potential sources in the environment, and because they tend to "sink," Group 2 VOCs are more likely to be found in ground waters than in surface waters. Like Group 1 VOCs, these compounds may be more likely to co-occur with each other than with other regulated VOCs, not because of their common potential source, but because they are part of the same degradation series. Co-occurrence in this case may be the occurrence of two or more distinct compounds in the same location but at different times (i.e., spatial but not temporal co-occurrence). For example, PCE might be found in ground water taken from a well subject to local PCE contamination. Over time, the level of PCE might diminish or disappear altogether, while the concentrations of TCE, the dichloroethylenes, and vinyl chloride, which might have been previously undetected, could rise due to microbial degradation.

There are other VOCs which share potential environmental sources, and thus might also co-occur. For example, 1,2-dichloroethane is a common additive to leaded gas, and so it might co-occur with the Group 1 VOCs if the major contaminant source contained leaded gasoline. Additionally, 1,1,1-trichloroethane may be transformed, via hydrolysis or microbial degradation, to 1,1,-dichloroethylene (Howard, 1997), which is part of the PCE degradation series. However, these associations are not as strong as the associations between the constituents of Group 1 and Group 2.

IV.B. Overview of Results

If there is a generalization that can be made about contaminant occurrence in drinking water systems, it is that occurrence is highly variable. It is important to note that all 64 contaminants regulated under the Phase rules (Table I.A.1) have been detected in drinking water systems. Only five of the 64 contaminants do *not* have a reported detection at >0.5 their MCL. They are all SOCs: glyphosate, hexachlorobenzene, methoxychlor, oxamyl, and 2,4,5-TP. All other IOCs, SOCs, and VOCs have detections >0.5 MCL.

The frequency of occurrence in water systems ranges widely. All IOCs, except asbestos, have been detected in 90 - 100% of systems in some States. Only three IOCs (antimony, beryllium, and cyanide) have not been detected in at least one system in at least one State (in either surface or ground water). The high occurrence rates are not surprising for IOCs because they occur naturally. Occurrence and concentration of IOCs can vary greatly related to natural geologic sources (sometimes enhanced by mineral extraction activities). Often such sources, and effects, are localized and are beyond the scope of this analysis.

Twenty-nine of the 30 regulated SOCs have not been detected at all in some States; only phthalate occurs in every State, but its high occurrence relates, in part, to sample contamination from plastics used in sampling or lab work (or system plumbing). Because of this problem, phthalate is not included among some of the other summary values presented. (Dioxin is not included among these summary values either, because dioxin only rarely has been included in monitoring.) The greatest range for any of the organic chemicals is for atrazine, which ranges from zero systems in some States to >97% occurrence in surface water systems in Midwestern States (up to 9% in ground water systems).

Fourteen of the 21 VOCs have not been detected in some of the participating States. The seven that occur in every State studied, in either surface or ground water systems, are: ethylbenzene, cis-1,2-dichloroethane, tetrachloroethylene (PCE), trichloroethylene (TCE), vinyl chloride, 1,1,1-trichlorethane, and xylenes. Many VOCs occur in up to, or more than, 30% of surface or ground water systems in various States.

Only nine of the 64 contaminants occur in fewer than 1% of either surface water or ground water systems in all States. These are all SOCs: alachlor, carbofuran, chlordane, glyphosate, hexachlorobenzene, hexachloropentadiene, oxamyl, toxaphene, and PCBs (though PCBs have not been monitored as intensively as the other SOCs). In contrast, the five compounds which occur most frequently at concentrations greater than the value of their MCL (which does not in itself constitute an MCL violation) are shown in Table IV.B.1. (As noted, phthalates and dioxin are excluded.)

Several observations are warranted. First, MCL exceedances are not common, even for this top-five list. The compounds with the most common exceedances are a mix of VOCs and SOCs. Fluoride makes the ground water list, but fluoride detections tend to be a result of natural occurrence and treatment additions. Three of the compounds occur both on the surface water and ground water list, suggesting their widespread and common impact. The data in Table IV.B.1 illustrate a general trend: exceedances for SOCs tend to be more common for surface water systems; for VOCs, exceedances tend to be more common for ground water systems (though the differences from surface water are not great; see Table IV.C.1-3, also). (It should be noted, however, that the detection of methylene chloride and EDB can be confounded with other organic compounds, and it is suspected that the occurrence values reported may be too great.)

Table IV.B.1. Five Phase rule contaminants which occur most frequently at concentrations greater than their MCL in either surface water or ground water systems. The five contaminants are those with the greatest percentage of systems showing exceedances in the national cross-section.

Surface-Water Supplied Systems		**Ground-Water Supplied Systems**	
Contaminant	**% Systems with MCL** exceedances**	**Contaminant**	**% Systems with MCL** exceedances**
atrazine	10.7%	DBCP	2.0%
EDB	3.7%	methylene chloride	2.3%
methylene chloride	4.7%	PCE*	1.8%
PCE*	1.7%	TCE*	1.5%
TCE*	1.2%	1,1,1-trichloro- ethane; and fluoride	1.3%

*PCE- tetrachloroethylene; TCE- trichloroethylene

** % Systems with MCL exceedances indicates the proportion of systems with any analytical results exceeding the concentration value of the MCL; it does not necessarily indicate an MCL violation. An MCL violation occurs when the MCL is exceeded by the average results from four quarterly samples or confirmation samples as required by the primacy State.

IV.C. Contaminant Occurrence in Surface Water and Ground Water Systems

Developing a further understanding of the differences in occurrence patterns, and the susceptibility or vulnerability of a water supply to contamination, is a primary objective of this work. Understanding the vulnerability of water supplies and water systems—the presence or probability of contaminant occurrence—is a key to developing cost-effective monitoring programs. Recognition of the differences between surface water and ground water systems has always been a basic element in developing monitoring strategies and protecting public health. Surface water systems are, in essence, vulnerable to any contaminant with a source in the upstream watershed, but they have the potential factor of dilution working for them to dampen impacts. Ground water sources are typically vulnerable to a smaller number of sources (those on flow paths leading to the supply wells) but have less capacity for dilution. As noted above, there are basic differences in occurrence patterns between these water sources, but these differences vary by contaminant. For example, 14% of the regulated SOCs are detected in more than 5% of surface water systems in the national cross-section. In contrast, no SOCs occur in more than 5% of ground water systems. Nearly two-thirds of SOCs are detected in >1% of surface water systems, but over three-fourths of these SOCs are detected in only a very few (<1%) ground water systems. For VOCs, the occurrence patterns reverse to some degree, but VOCs are more common than most SOCs in both surface and ground water systems. VOCs are detected in surface water systems more frequently than is generally recognized.

Tables IV.C.1 through IV.C.3 summarize occurrence patterns for select Phase II/V contaminants and the contaminant groups. The individual contaminants represent a range of occurrence patterns within each contaminant group, e.g., SOCs ranging from atrazine to glyphosate. Tables IV.C.1 and 2 summarize the national cross-section occurrence values, with ranges shown from all State data, for surface and ground water. Table IV.C.3 shows a comparison of national cross-section values for surface and ground water. IOCs are about equally common (detected >MRL) in surface water systems and ground water systems, but ground water systems have higher concentrations, i.e., a greater proportion of systems exceeding 0.5 MCL and the MCL. SOCs are far more common in surface water systems than ground water systems and surface water systems have more exceedances as well. (The Group 2 SOCs, which have more equal occurrence in surface water systems and ground water systems, include the phthalates and their attendant problems.) VOCs show much greater general occurrence in surface water systems but surface water systems and ground water systems are nearly equal for the occurrence of exceedances. Whether a greater proportion of surface water systems or ground water systems show detections varies by individual VOC or group. In total, ground water systems show slightly more MCL violations for VOCs (Table IV.C.4).

Table IV.C.1. Summary of occurrence of selected Phase II/V contaminants in water systems using surface water, from national cross-section States; ranges from all States studied. Percent MCL violations derived from SDWIS information for 1/1/93-3/31/1998.

Contaminant	**Percent > MRL**		**Percent > ½ MCL**		**Percent > MCL***		**National Percent MCL Violations**
	Cross-Section	**Range (All States)**	**Cross-Section**	**Range (All States)**	**Cross-Section**	**Range (All States)**	
Individual Contaminants (in surface water systems)							
IOCs							
Barium	49.1%	22.1% - 100.0%	0.6%	0.0% - 6.1%	0.5%	0.0% - 3.0%	0.0%
Cadmium	5.1%	0.0% - 100.0%	1.3%	0.0% - 9.1%	0.2%	0.0% - 1.9%	<0.1%
Mercury	9.0%	0.0% - 100.0%	1.3%	0.0% - 6.1%	0.5%	0.0% - 1.9%	<0.1%
SOCs							
2,4-D	11.2%	0.0% - 50.5%	0.2%	0.0% - 0.6%	0.0%	0.0% - 0.0%	0.0%
Atrazine	21.1%	0.0% - 97.1%	13.2%	0.0% - 62.4%	10.7%	0.0% - 51.4%	0.8%
Benzo[a]pyrene	0.5%	0.0% - 3.7%	0.0%	0.0% - 1.9%	0.0%	0.0% - 0.0%	0.0%
Carbofuran (Furadan)	0.8%	0.0% - 1.8%	0.0%	0.0% - 0.0%	0.0%	0.0% - 0.0%	0.0%
Ethylene Dibromide (EDB)	4.2%	0.0% - 9.4%	3.8%	0.0% - 8.5%	3.7%	0.0% - 8.5%	0.1%
Glyphosate (Roundup)	0.0%	0.0% - 0.9%	0.0%	0.0% - 0.0%	0.0%	0.0% - 0.0%	0.0%
Simazine	15.9%	0.0% - 67.0%	2.5%	0.0% - 11.9%	1.0%	0.0% - 4.6%	0.0%
VOCs							
Benzene	3.9%	0.0% - 17.1%	0.5%	0.0% - 3.0%	0.3%	0.0% - 3.0%	<0.1%
Styrene	4.1%	0.0% - 14.5%	0.0%	0.0% - 0.0%	0.0%	0.0% - 0.0%	0.0%
Tetrachloroethylene	7.1%	0.0% - 16.2%	2.5%	0.0% - 15.2%	1.7%	0.0% - 15.2%	0.1%
Vinyl Chloride	3.1%	0.0% - 13.7%	0.3%	0.0% - 3.0%	0.3%	0.0% - 3.0%	0.0%
Xylenes (Total)	12.3%	1.9% - 29.1%	0.2%	0.0% - 1.9%	0.0%	0.0% - 0.0%	0.0%
Group Summaries							
IOCs							
IOCs-Regulated	83.7%	69.7% - 100.0%	6.8%	0.0% - 92.6%	2.5%	0.0% - 6.1%	0.3%
SOCs							1.0%
SOCs-Group 1	21.9%	0.0% - 97.1%	13.2%	0.0% - 62.4%	10.7%	0.0% - 51.4%	
SOCs-Group 2	20.4%	0.0% - 88.9%	1.2%	0.0% - 5.6%	0.9%	0.0% - 2.8%	
VOCs							
VOCs-Regulated	41.1%	8.3% - 89.7%	15.4%	0.0% - 60.6%	8.2%	0.0% - 57.6%	0.2%
VOCs-Group 1	19.5%	2.1% - 42.9%	1.7%	0.0% - 12.1%	0.9%	0.0% - 6.1%	
VOCs-Group 2	11.1%	0.0% - 30.3%	4.5%	0.0% - 30.3%	2.9%	0.0% - 30.3%	

* % > MCL indicates the proportion of systems with any analytical results exceeding the concentration value of the MCL; it does not necessarily indicate an MCL violation. An MCL violation occurs when the MCL is exceeded by the average results from four quarterly samples or confirmation samples as required by the primacy State.

IOC-Regulated: includes all the regulated IOCs.
SOCs-Group 1: includes alachlor, atrazine, and simazine.
SOCs-Group 2: includes bis(2-ethylhexyl)phthalate, bis(2-ethylhexyl)adipate, and benzo(a)pyrene.
VOCs-Regulated: includes all the regulated VOCs
VOCs-Group 1: includes benzene, ethyl benzene, toluene, and total xylenes (LNAPLs).
VOCs-Group 2: includes cis-1,2-dichloroethylene, trans-1,2-dichloroethylene, 1,1-dichloroethene, tetrachloroethylene, trichloroethene, and vinyl chloride (DNAPLs).

Table IV.C.2. Summary of occurrence of selected Phase II/V contaminants in water systems using ground water, from national cross-section States; ranges from all States studied. Percent MCL violations derived from SDWIS information for 1/1/93-3/31/1998.

Contaminant	Percent > MRL		Percent > ½ MCL		Percent > MCL*		National Percent MCL Violations
	Cross-Section	**Range (All States)**	**Cross-Section**	**Range (All States)**	**Cross-Section**	**Range (All States)**	
Individual Contaminants (in ground water systems)							
IOCs							
Barium	47.3%	31.5% - 98.6%	0.8%	0.0% - 6.9%	0.2%	0.0% - 3.1%	0.1%
Cadmium	4.9%	0.4% - 93.9%	1.2%	0.0% - 21.7%	0.6%	0.0% - 1.5%	0.1%
Mercury	4.5%	0.4% - 93.3%	0.7%	0.0% - 3.2%	0.4%	0.0% - 2.0%	<0.1%
SOCs							
2,4-D	1.2%	0.0% - 4.7%	0.0%	0.0% - 0.1%	0.0%	0.0% - 0.1%	<0.1%
Atrazine	2.0%	0.0% - 9.2%	0.3%	0.0% - 2.0%	0.1%	0.0% - 0.5%	<0.1%
Benzo[a]pyrene	0.5%	0.0% - 2.9%	0.1%	0.0% - 0.3%	0.1%	0.0% - 0.3%	0.0%
Carbofuran (Furadan)	0.1%	0.0% - 0.1%	0.0%	0.0% - 0.1%	0.0%	0.0% - 0.0%	0.0%
Ethylene Dibromide (EDB)	1.0%	0.0% - 2.6%	1.0%	0.0% - 4.2%	0.7%	0.0% - 1.9%	<0.1%
Glyphosate (Roundup)	0.1%	0.0% - 0.5%	0.0%	0.0% - 0.0%	0.0%	0.0% - 0.0%	0.0%
Simazine	1.4%	0.0% - 3.2%	0.0%	0.0% - 0.1%	0.0%	0.0% - 0.0%	0.0%
VOCs							
Benzene	1.2%	0.0% - 4.2%	0.5%	0.0% - 1.8%	0.4%	0.0% - 1.8%	<0.1%
Styrene	2.1%	0.0% - 8.1%	0.2%	0.0% - 0.9%	0.2%	0.0% - 0.9%	0.0%
Tetrachloroethylene	4.3%	0.6% - 14.2%	2.3%	0.1% - 7.6%	1.8%	0.0% - 5.7%	0.1%
Vinyl Chloride	0.5%	0.0% - 3.0%	0.2%	0.0% - 0.6%	0.2%	0.0% - 0.6%	<0.1%
Xylenes (Total)	3.9%	0.4% - 15.3%	0.2%	0.0% - 1.0%	0.1%	0.0% - 0.8%	0.0%
Group Summaries							
IOCs							
IOCs-Regulated	83.5%	73.5% - 100.0%	9.2%	0.0% - 58.9%	4.2%	0.0% - 6.2%	0.6%
SOCs							0.1%
SOCs-Group 1	2.4%	0.0% - 9.2%	0.3%	0.0% - 2.0%	0.1%	0.0% - 0.6%	
SOCs-Group 2	13.4%	0.0% - 54.6%	1.9%	0.0% - 9.1%	1.0%	0.0% - 9.1%	
VOCs							
VOCs-Regulated	19.9%	6.2% - 65.5%	7.9%	1.3% - 28.7%	6.1%	0.7% - 27.9%	0.4%
VOCs-Group 1	6.6%	2.0% - 28.6%	1.3%	0.0% - 6.2%	0.9%	0.0% - 4.4%	
VOCs-Group 2	6.4%	2.1% - 21.1%	3.9%	0.9% - 12.9%	3.2%	0.0% - 12.9%	

* % > MCL indicates the proportion of systems with any analytical results exceeding the concentration value of the MCL; it does not necessarily indicate an MCL violation. An MCL violation occurs when the MCL is exceeded by the average results from four quarterly samples or confirmation samples as required by the primacy State.

IOC-Regulated: includes all the regulated IOCs.
SOCs-Group 1: includes alachlor, atrazine, and simazine.
SOCs-Group 2: includes bis(2-ethylhexyl)phthalate, bis(2-ethylhexyl)adipate, and benzo(a)pyrene.
VOCs-Regulated: includes all the regulated VOCs
VOCs-Group 1: includes benzene, ethyl benzene, toluene, and total xylenes (LNAPLs).
VOCs-Group 2: includes cis-1,2-dichloroethylene, trans-1,2-dichloroethylene, 1,1-dichloroethene, tetrachloroethylene, trichloroethene, and vinyl chloride (DNAPLs).

Table IV.C.3. Summary comparison of occurrence of selected Phase II/V contaminants in water systems using surface water vs. ground water, from national cross-section States. Percent MCL violations derived from SDWIS information for 1/1/93-3/31/1998.

Contaminant	**Percent > MRL**		**Percent > ½ MCL**		**Percent > MCL***		**National % MCL Viol.-SW**	**National % MCL Viol.-GW**
	Surface Water	**Ground Water**	**Surface Water**	**Ground Water**	**Surface Water**	**Ground Water**		
Individual Contaminants								
IOCs								
Barium	49.1%	47.3%	0.6%	0.8%	0.5%	0.2%	0.0%	0.1%
Cadmium	5.1%	4.9%	1.3%	1.2%	0.2%	0.6%	<0.1%	0.1%
Mercury	9.0%	4.5%	1.3%	0.7%	0.5%	0.4%	<0.1%	<0.1%
SOCs								
2,4-D	11.2%	1.2%	0.2%	0.0%	0.0%	0.0%	0.0%	<0.1%
Atrazine	21.1%	2.0%	13.2%	0.3%	10.7%	0.1%	0.8%	<0.1%
Benzo[a]pyrene	0.5%	0.5%	0.0%	0.1%	0.0%	0.1%	0.0%	0.0%
Carbofuran (Furadan)	0.8%	0.1%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Ethylene Dibromide (EDB)	4.2%	1.0%	3.8%	1.0%	3.7%	0.7%	0.1%	<0.1%
Glyphosate (Roundup)	0.0%	0.1%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Simazine	15.9%	1.4%	2.5%	0.0%	1.0%	0.0%	0.0%	0.0%
VOCs								
Benzene	3.9%	1.2%	0.5%	0.5%	0.3%	0.4%	<0.1%	<0.1%
Styrene	4.1%	2.1%	0.0%	0.2%	0.0%	0.2%	0.0%	0.0%
Tetrachloroethylene	7.1%	4.3%	2.5%	2.3%	1.7%	1.8%	0.1%	0.1%
Vinyl Chloride	3.1%	0.5%	0.3%	0.2%	0.3%	0.2%	0.0%	<0.1%
Xylenes (Total)	12.3%	3.9%	0.2%	0.2%	0.0%	0.1%	0.0%	0.0%
Group Summaries								
IOCs								
IOCs-Regulated	83.7%	83.5%	6.8%	9.2%	2.5%	4.2%	0.3%	0.6%
SOCs							1.0%	0.1%
SOCs-Group 1	21.9%	2.4%	13.2%	0.3%	10.7%	0.1%		
SOCs-Group 2	20.4%	13.4%	1.2%	1.9%	0.9%	1.0%		
VOCs								
VOCs-Regulated	41.1%	19.9%	15.4%	7.9%	8.2%	6.1%	0.2%	0.4%
VOCs-Group 1	19.5%	6.6%	1.7%	1.3%	0.9%	0.9%		
VOCs-Group 2	11.1%	6.4%	4.5%	3.9%	2.9%	3.2%		

* % > MCL indicates the proportion of systems with any analytical results exceeding the concentration value of the MCL; it does not necessarily indicate an MCL violation. An MCL violation occurs when the MCL is exceeded by the average results from four quarterly samples or confirmation samples as required by the primacy State.

IOC-Regulated: includes all the regulated IOCs.
SOCs-Group 1: includes alachlor, atrazine, and simazine.
SOCs-Group 2: includes bis(2-ethylhexyl)phthalate, bis(2-ethylhexyl)adipate, and benzo(a)pyrene.
VOCs-Regulated: includes all the regulated VOCs
VOCs-Group 1: includes benzene, ethyl benzene, toluene, and total xylenes (LNAPLs).
VOCs-Group 2: includes cis-1,2-dichloroethylene, trans-1,2-dichloroethylene, 1,1-dichloroethene, tetrachloroethylene, trichloroethene, and vinyl chloride (DNAPLs).

Table IV.C.4. Summary of national MCL violations for surface water and ground water systems. Percent MCL violations derived from SDWIS information for 1/1/93-3/31/98.

Contaminant Group	**Percentage of systems with MCL violations**	
	Surface Water Systems	**Ground Water Systems**
IOCs	0.3%	0.6%
SOCs	1.0%	0.1%
VOCs	0.2%	0.4%

In short, surface water systems tend to be more vulnerable than ground water systems for many contaminants, but they are not necessarily more vulnerable to occurrences exceeding an MCL. A greater percentage of surface water systems show detections of SOCs than ground water systems, in particular, but surface water systems show greater occurrence of VOCs as well. Surface water systems even show a larger proportion of systems with exceedances of VOC-MCL concentrations, though ground water systems show a slightly larger proportion of systems with MCL violations. Partly, this may be related to the greater temporal variability in contaminant occurrence and transport in surface waters. Surface water is vulnerable to many problems related to land surface discharges of contaminants through runoff, spills, and even shallow ground water recharge. Yet even in the areas most vulnerable to SOC contamination, variability is still a key and confounding factor. In Midwest States, where 90-97% of surface water systems show occurrence of pesticides, such as atrazine, there is still the small percentage of systems that have no detections. These are often small watersheds (reservoirs) that are protected from contaminant sources.

IV.C.1. Ground Water Vulnerability

Ground water systems are much more variable in nature than surface water systems because ground water adds other dimensions to the vulnerability equation. Aquifer type and well depth are key hydrogeologic factors affecting ground water vulnerability. Confined aquifers have a mantle of material that inhibits the recharge of water with modern contaminants into an aquifer. Depth of a well and the resultant depth of the ground water influences vulnerability. Depth, in a general sense, can be related to the time it takes for recharge water to reach a well intake in an aquifer. The water tapped by a deep well may simply be too old to contain any modern, synthetic organic contaminants. Deeper wells are prone to occurrence of natural contaminants, such as some IOCs or radiochemicals, but are less susceptible to occurrence of anthropogenic contaminants such as pesticides or many VOCs.

Aggregating all ground water systems in a vulnerability assessment does not provide adequate information. Most public water system databases contain little information that can be used to further unravel ground water system vulnerability. Further complicating such studies, is that many public water systems operate multiple wells; some of which may be shallow and vulnerable to contaminants, others may be deep and protected. When analyzing finished water data it is often not possible to discern the ultimate source of a sample. Vulnerability is best defined using water quality data from a particular well, where aquifer, depth, and other factors can be evaluated independently. To illustrate, data from Iowa will be used where investigators have matched basic aquifer and well properties to SDWA compliance monitoring data (Hallberg et al., 1996).

In Iowa, the Iowa Department of Natural Resources and the U.S. Geological Survey have sampled municipal water-supply wells for common contaminants (Kolpin et al., 1997). These data show that wells less than 100 feet deep have greater detections of nitrate and some common herbicides, for example (Table IV.C.5). These studies also show that there are differences in occurrence rates among different aquifer types; alluvial aquifers having more common detections than bedrock aquifers, with glacial-drift aquifers also exhibiting lower vulnerability. This trend in aquifer type reflects two inter-related variables: depth and confinement. Alluvial aquifers are very shallow, by definition, and typically have a thin confining cover, if any. The glacial-drift aquifers have a variable thickness of glacial deposits that are confining beds, and tend to be relatively deeply buried. Bedrock aquifers are quite variable in setting and depth. The data are for the raw water drawn from the wells. Although most of these ground water systems would not have treated water in a way that would affect the occurrence of these contaminants, they may have blended water that would lower the concentration and/or the occurrence.

Table IV.C.5. Percentage of Iowa municipal water-supply wells with detections of various contaminants: for wells of various depths and for wells finished in different aquifers. (After Kolpin, Sneck-Fahrer, Hallberg, and Libra, 1997.)

	Percent PWS wells with detections of:				
Systems using:	**nitrate**	**alachlor**	**atrazine**	**cyanazine**	**metolachlor**
Wells <50 feet deep	76%	4%	16%	4%	7%
Wells 51-100 feet deep	76%	1%	22%	9%	29%
Wells >100 feet deep	28%	2%	8%	0%	3%
Systems using wells finished in:					
Alluvial Aquifers	74%	3%	21%	7%	17%
Bedrock Aquifers	46%	2%	15%	0%	5%
Glacial-Drift Aquifers	35%	0%	0%	0%	0%

Iowa studies have also linked SDWA compliance monitoring data from finished drinking water with various well and aquifer characteristics. These water-quality data come from standard drinking water samples (entry points to distribution systems). Similar occurrence patterns can be seen in these data (Table IV.C.6). First, a greater occurrence of contaminants in surface water systems than ground water systems is pronounced, except for the Group 2 VOCs, (the dense VOCs: PCE, etc.), which are more prevalent in ground water systems. Within the systems using ground water, those using alluvial aquifers generally show the highest occurrence of SOCs and VOCs, followed closely by the vulnerable karst aquifers. The exception, VOCs-Group 1, the "light" petroleum-related compounds, are essentially equal in the various aquifers.

These data illustrate the problems with analyzing ground water as a single entity. For example, overall, 9% of ground water systems have detections of atrazine, but 30% of systems using alluvial aquifer wells less than 50 feet deep show detections. Within each aquifer type, there is a general decrease in contaminant occurrence as well depth increases. Ground water systems using alluvial aquifers show the most pronounced change with well-depth.

In the bedrock aquifers, contaminants are evident even in systems with wells over 250 feet deep. There are two factors that contribute to this. First, many of these systems use multiple wells from different sources. Many have a deeper well (>250 feet) and a shallower well (<150 feet). The finished water quality data from the composite system confounds this analysis because the contaminant occurrence cannot be uniquely related to wells of a given depth. Some of these systems do show contaminant occurrence at greater depth, particularly in the karst aquifer group. The hydrogeologic conditions where these aquifers are prevalent promote more rapid ground water circulation, to greater depth, than in most areas. Hence, younger water, with SOCs, VOCs, or other human-induced contamination, extends to greater depth than in many areas. The combination of aquifer characteristics, hydrogeologic setting, and well depth must all be factored to understand some of these patterns.

Table IV.C.6. Percentage of Iowa public water systems with detections of various contaminants or contaminant groups, for systems using different source water, aquifers, and wells of different depths. (Hallberg, Riley, Kantamneni, Weyer, and Kelley, 1996.)

	Percent systems with detections of:				
Systems using:	**atrazine**	**Any SOCs**	**Any VOCs**	**VOCs-1 ("light")**	**VOCs-2 ("heavy")**
All Surface Water	97%	97%	37%	23%	3%
All Ground Water	9%	10%	16%	9%	4%
Alluvial Aquifer (all)	23%	29%	23%	13%	9%
With Wells < 50 ft deep	30%	33%	26%	14%	9%
51-100 feet	10%	18%	18%	13%	8%
101-150 feet	0%	10%	7%	0%	7%
Fractured-Karst Bedrock Aquifer (a) (all)	14%	14%	21%	14%	6%
With Wells < 150 ft deep	22%	22%	87%	16%	9%
151-250 feet	24%	24%	22%	15%	10%
>250 feet	7%	7%	15%	13%	3%
Sandstone Bedrock Aquifer (b) (all)	2%	3%	20%	14%	4%
With Wells < 150 ft deep	2%	2%	50%	50%	40%
151-250 feet	11%	11%	33%	22%	0%
>250 feet	1%	2%	18%	13%	3%

(a) Silurian-Devonian Aquifer; (b) Cambrian-Ordovician Aquifer

Figure IV.C.1 provides a graphic illustration, showing the occurrence of nitrate in wells of various depths in two regions of Iowa. These data are from private and public wells, but help to illustrate these points. In the north-central region (NC), there is a pronounced relationship between nitrate concentration and depth; nitrate levels decline sharply in wells >100 feet deep. The karst aquifer hydrogeologic settings, described above, are from the northeastern region (NE). Here the hydrogeologic setting promotes deeper circulation of ground water, and the nitrate-depth relationship, while still apparent, is not as sharply defined. Nitrate occurs at high concentrations to greater depth in this environment. The graphs also illustrate the variability in ground water systems; while high concentrations of contaminants may be more prevalent at shallower depths, there is still a mode of low occurrence, because of the range and variability of hydrogeologic controls.

Given the variability in ground water settings, and system level variations, developing improved monitoring requires refined information at the sub-state and system levels. Information on aquifer characteristics, well depth, well casing and construction details, wellhead protection information, as well as system and well maintenance history are some of the factors that should be considered. Several States have utilized such information to improve the cost-effectiveness of their monitoring program. The NJ Department of Environmental Protection, after conducting a study of drinking water vulnerability with the USGS, estimated that waiving monitoring for wells and intakes that are not vulnerable to pesticide contamination will save an estimated $5.1 million dollars per year. The Washington State Department of Health, which also conducted a study with the USGS, estimated that pesticide monitoring waivers for public supply wells would save an estimated $6.0 million dollars per year (Vowinkel et al., 1996; Ryker and Williamson, 1996).

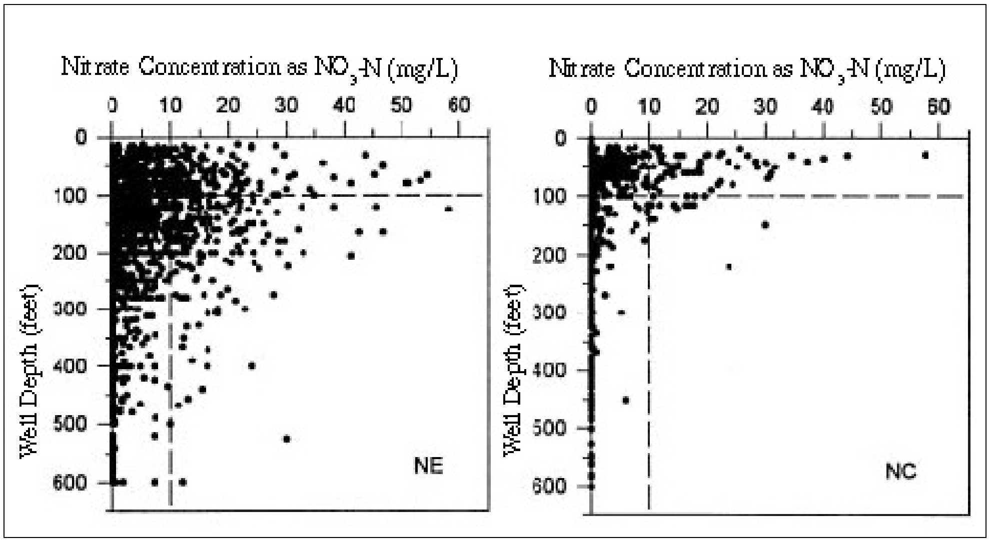


Figure IV.C.1. Nitrate-N concentrations versus well depth, from water-quality analyses of private wells in north-central (NC) and northeastern (NE) Iowa, illustrating the general inverse relationship between well depth (or ground-water depth) and contaminant occurrence. (Kross et al., 1990, 1993; Hallberg et al., 1992).

IV.D. National Cross-Section: Perspectives and Comparison with Other Data

Before proceeding with further analysis of the State SDWA data, some perspective on the value of the national cross-section can be offered. As noted, the occurrence estimates developed from the cross-section cannot be considered a truly "representative" sample of nation-wide occurrence of a given contaminant. If the States included do represent an adequate cross-section, these values should be indicative of the average values. With the relatively low occurrence rates and variability apparent in the data, the cross-section provides what is most likely a high estimate of contaminant occurrence. There are few other sources of data available to use as a check on this sample. Three primary sources of data will be used to provide perspective: (1) the URCIS database; (2) the atrazine and simazine studies by Novartis; and (3) data from the USGS.

IV.D.1. URCIS Data

The URCIS database, as described above, is one of the few national compilations of drinking water contaminant data. While URCIS has data quality limitations, it also has greater State representation (some data from 39 States and 1 territory) than this study. As noted, for States that had relatively complete data in both URCIS and this study, data comparisons were made and the differences were found to be minor. (The URCIS database is a subset of the data used for this study.) Some of the contaminants that were monitored under the original unregulated contaminant monitoring regime (i.e., the URCIS data) have since become regulated contaminants; hence, there is more data in the current State databases than in URCIS. Comparing national occurrence values computed from URCIS with the national cross-section shows that, in the majority of cases (about 67%), the cross-section has a slightly higher proportion of systems with detections of a contaminant with overlapping data (Table IV.D.1), and for about one-third of contaminants, URCIS shows an equal or greater percentage. In all cases, the values are comparable; no values for systems with detections stand out as markedly or unexpectedly different. In each case, contaminants that stand out as exceptionally high or low, are proportionally similar in the other data set (e.g., methylene chloride, xylenes, or p-chlorotoluene). This suggests that if additional State data were added, in a representative fashion, that many of the national cross-section values would decline slightly. An EPA survey for VOCs conducted in the 1980s showed similar results also, in terms of the proportions of systems with VOCs and the general order of occurrence (Westrick et al., 1984).

IV.D.2. Novartis Atrazine and Simazine Data

Novartis Crop Protection, Inc. conducted a special regional analysis of atrazine and simazine occurrence in CWSs, as part of a special study under direction from EPA's Office of Pesticide Programs. While limited to two contaminants, this study is important because it compiled data from CWS monitoring, conducted from 1993 through 1996, from 21 States across the country, for one of the most widely occurring compounds (Clarkson et al., 1997). It represents one of the most complete assessments of drinking water occurrence data of its kind. The 21 States selected were the States with the greatest use for these pesticides, so occurrence estimates from this study should be high (relative to occurrence for all 50 States, for example). The Novartis study includes data from several States in the cross-section (California, Illinois and Michigan) and from other State data reviewed and used in the ranges presented in this study (Indiana, Iowa, Ohio). The individual State estimates matched quite closely. A comparison of the national estimates (Table IV.D.2.) shows that the cross-section provides a lower estimate for atrazine and a very comparable estimate for simazine. The ranges are nearly identical, which suggests that data reviewed for this study capture a similar range of conditions as the Novartis sample of 21 States.

Table IV.D.1. Comparison of national occurrence estimates from URCIS and the national cross-section of State SDWA data.

	SURFACE WATER SYSTEMS			**GROUND WATER SYSTEMS**		
	URCIS	**National Cross-Section**		**URCIS**	**National Cross-Section**	
	%>MRL	**%>MRL**	**RANGE**	**%>MRL**	**%>MRL**	**RANGE**
SOCs - Regulated						
Ethylene Dibromide (EDB)	2.5%	4.1%	0%-9.4%	1.1%	1.2%	0%-4.0
Dibromochloropropane (DBCP)	2.6%	5.3%	0%-13.2%	2.5%	3.0%	0%-9.3%
VOCs - Regulated						
Benzene	2.5%	4.2%	0%-17.1%	1.9%	1.3%	0.5%-4.2%
Methylene chloride	11.9%	24.3%	0%-73.5%	3.8%	9.6%	1.8%-57.8%
Styrene	1.4%	4.5%	0%-14.5%	0.6%	2.3%	0%-8.1%
Tetrachloroethylene	3.1%	7.6%	0%-16.2%	4.2%	4.4%	0%-14.2%
Toluene	8.2%	10.8%	0%-30.0%	3.8%	4.0%	1.6%-14.4%
Trichloroethylene	6.7%	5.4%	0%-21.2%	4.2%	2.9%	0.6%-12.8%
Vinyl chloride	1.2%	3.4%	0%-13.7%	0.5%	0.5%	0%-3.0%
Xylenes (Total)	11.9%	13.3%	0%-29.1%	3.1%	4.2%	0.9%-15.3%
VOCs - Unregulated						
(m-) 1,3-Dichlorobenzene	0.8%	0.8%	0%-2.9%	0.2%	0.2%	0%-1.3%
1,1,1,2-Tetrachloroethane	0.6%	1.0%	0%-6.1%	0.1%	0.2%	0%-0.7%
Chloromethane	2.4%	6.5%	0%-28.6%	1.5%	1.1%	0%-9.7%
1,2,3-Trichloropropane	0.3%	0.8%	0%-9.1%	0.3%	0.3%	0%-0.9%
Bromomethane	1.3%	1.0%	0%-3.0%	0.7%	0.3%	0%-2.3%
o-Chlorotoluene	0.6%	0.3%	0%-1.9%	0.2%	0.4%	0%-.9%
p-Chlorotoluene	0.3%	0.3%	0%-1.9%	0.2%	0.3%	0%-0.4%

Table IV.D.2. Comparison of national estimates of occurrence (percentage of water systems with detections) of atrazine and simazine from the Novartis study of 21 high-use States and from the national cross-section (8 States) compiled in this study.

	Percentage of public water systems with:			
	Atrazine		**Simazine**	
	% >MRL	**Range**	**% >MRL**	**Range**
National cross-section surface water systems	21.1%	0-97.1%	15.9%	0-67%
Novartis surface water systems	51.7%	0-98.4%	15.1%	0-69%
National cross-section ground water systems	2.0%	0-9.2%	1.4%	0-3.2%
Novartis ground water systems	3.0%	0-15.0%	0.6%	0-3.8%

IV.D.3. USGS Data

A third source of national data are various studies by the USGS, particularly their National Water Quality Assessment Program (NAWQA). Results from USGS studies have been brought up repeatedly in CMR deliberations, noting that USGS estimates of pesticide ocurrences in surface and ground water are greater than EPA estimates. The USGS conducts various monitoring activities under nationally consistent protocols. This makes the data very valuable for reviewing national patterns of occurrence that can provide perspectives on drinking water problems. One important difference, however, is that the majority of the USGS data are from ambient water (i.e., raw water) and *not* drinking water. EPA estimates, such as in this study, are for drinking water (i.e., finished water, after treatment, that may have been blended from multiple raw water sources). While this precludes direct comparisons, the USGS work still provides some pertinent perspectives. The review of the USGS findings provide a comparitive national perspective and also illustrate why conclusions from the USGS ambient data are inherently different than drinking water analyzed by EPA.

In their review of the NAWQA data collected to date, Kolpin et al. (1998) note that detections of pesticide compounds are widespread in the shallow ground water sampled throughout the U.S., with one or more pesticide compounds detected in over 54% of the wells and springs sampled. Thirty-nine different pesticide compounds were detected; 26 at concentrations above 0.01 µg/L. This is a much higher frequency of occurrence than might be expected in the SDWA data. But further scrutiny is required to put these values in perspective. The USGS is analyzing for a broad spectrum of compounds, and the 54% value is for their entire contaminant group. Only five of the compounds are regulated under SDWA Phase II/V rules. Five more are on the current unregulated contaminant list that is required for monitoring by PWSs; twelve more are on the list for monitoring under the proposed new Unregulated Contaminant Monitoring Regulation.

Consistent with findings from previous large scale studies, more than 95% of the pesticide detections were at concentrations less than 1.0 µg/L. This concentration is below most of the MCLs for the regulated compounds; hence there are few exceedances. In contrast, the overall frequency of pesticide detection in the NAWQA studies is considerably higher than those reported by two previous large-scale, multi-state studies of pesticide occurrence in ground water across the United States: EPA's National Pesticide Survey (NPS) and National Alachlor Well Water Survey (NAWWS).

The USGS authors note that the contrasts in pesticide detection frequencies between the NAWQA data and the NPS and NAWWS investigations are attributable primarily to: "(a) low detection threshold characteristics of the analytical method used for NAWQA and (b) the relatively young age of ground water sampled in the NAWQA land-use studies." These are important factors for understanding the USGS information and putting it into perspective for drinking water programs: (1) USGS uses much lower minimum reporting levels than normal SDWA labs; (2) USGS studies are targeted to relatively vulnerable conditions to establish the link between ground water quality and land use, not drinking water and land use; and (3) the sampling is targeted at shallow wells, monitoring wells, and is not typically from deeper drinking water well (and thus the water sampled is relatively young).

The effects of targeting shallow, vulnerable ground water in drinking water quality studies were illustrated in the previous section. As the Iowa data show, 9% of all systems using ground water had atrazine detections, but 30% of systems using alluvial aquifer wells <50 feet deep had detections (Table IV.C.7). Even these drinking water supply wells would be somewhat deeper than many research monitoring wells.

The effect of the lower MRLs used in the USGS studies is to increase the number (and percentage) of detections. The USGS MRL is about 10-times lower than the MRL used in the EPA-NPS study (and most SDWA analyses), and the percentage of atrazine detections was nearly 40 times greater (Table IV.D.3). In a more direct comparison from a drinking water study, the State of Washington and the USGS cooperated on an evaluation of the vulnerability of PWS ground water systems to improve Washington's waiver system for required monitoring (Washington Div. of Drinking Water, 1995; Ryker and Williamson, 1996). Public water system wells were sampled and analyzed for various pesticides (and other SOCs). The special waiver study used lower MRLs than normally required for standard EPA or SDWA analyses in order to provide added quality assurance to their evaluation. This resulted in a substantial increase in the number of detections (Table IV.D.4). For example, of the 51 detections of atrazine with the special, lower MRL, only 17 would be detections using the normal SDWA MRLs.

Table IV.D.3. Percentage detections of atrazine in different occurrence studies and the minimum reporting levels used. (In part, from Kolpin et al., 1998.)

Study	**MRL-µg/L**	**% of Wells with Detections**
USGS NAWQA	0.01	28%
National Alachlor Well Water	0.03	12%
USEPA National Pesticide Survey	0.12	0.7%
Typical SDWA Analysis	0.1	

Table IV.D.4. Pesticide detections from public water system wells at different minimum reporting levels. From a special vulnerability and waiver study conducted by the State of Washington and the USGS (Washington Div. of Drinking Water, 1995; Ryker and Williamson, 1996).

Pesticide	**Number of detections with lower MRLs from special waiver study**	**Number of detections above standard EPA/SDWA MRLs**
2,4-D	6	6
2,4,5-TP	10	1
Atrazine	51	17
Lindane	1	0
Picloram	7	2
Simazine	17	2

While the USGS studies have important implications for many aspects of the drinking water program, the findings on occurrence in raw ambient water can not be directly related to drinking water.

In summary, the aggregated national cross-section of contaminant occurrence data compiled in this report appears to provide a conservative, but reasonable approximation of national occurrence values. The comparisons with other data suggest that the cross-section summary occurrence values are likely slightly high. Yet the data are representative enough that some basic principles can be derived to guide further data analysis and collection efforts.

V. CONTAMINANT OCCURRENCE: SYSTEM SIZE AND OTHER VULNERABILITY FACTORS

In addition to evaluating the fundamental differences and vulnerability of surface water and ground water supplied systems, other factors were evaluated to assess occurrence patterns and vulnerability. This section discusses system size, contaminant source issues, and temporal (seasonal) variability issues.

V.A. System Size and Contaminant Occurrence

Concerns are often raised about the greater problems of contaminant occurrence in small systems. In part, this is related to the more limited resources and options that small systems have when dealing with contaminant problems. Some questions raised during CMR deliberations indicate a perception that small systems in general may have greater contaminant occurrence. Some of the State data and all of the URCIS data were analyzed in relation to system size (by standard categories of the population served) and subdivided by water source (surface or ground) to assess whether any consistant trends are apparent in the data.

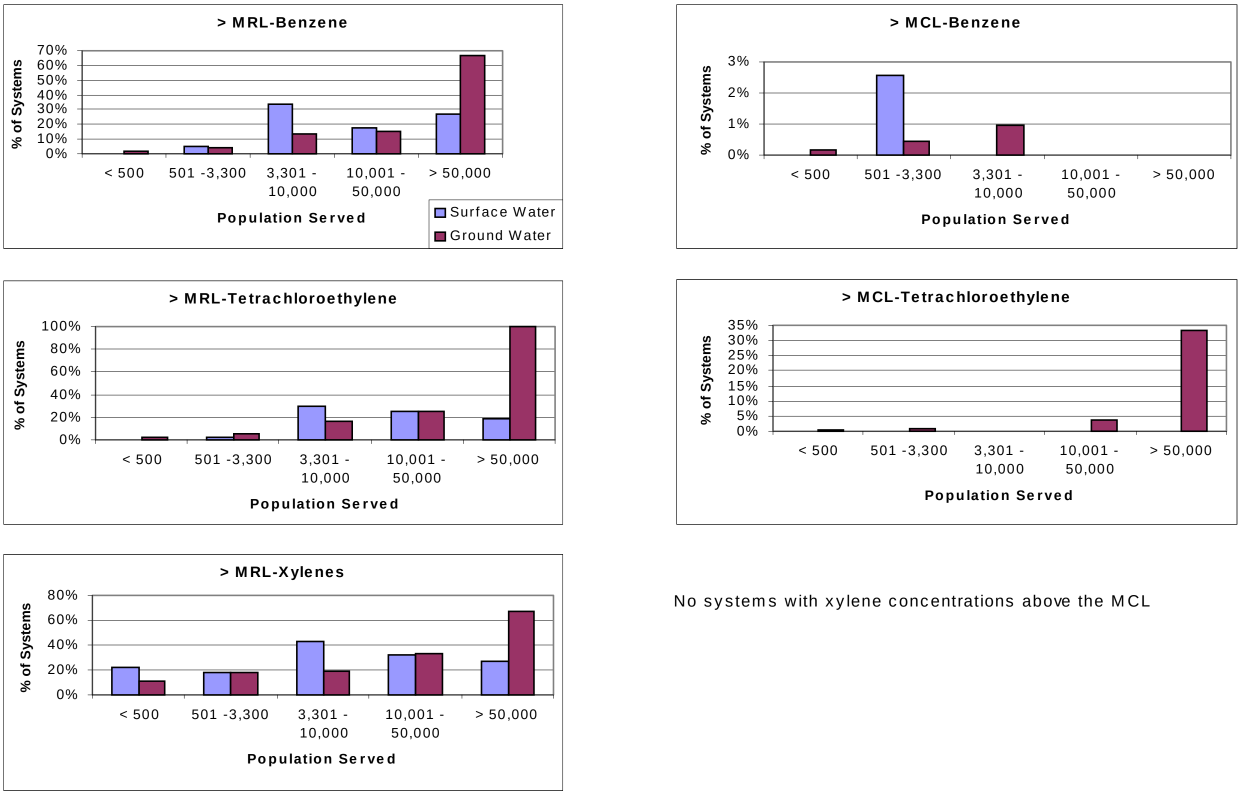
Surprisingly consistent patterns were revealed in all the data sets analyzed. Results for the selected Phase II/V contaminants (introduced in Tables IV.C.2-4) and various URCIS data are summarized in Tables V.A.1 through V.A.8 and in Figures V.A.1 through V.A.3 (Tables V.A.1 through V.A.8 are located in Appendix D). The Figures provide a summary of the trends in occurrence patterns. In general, the same consistent patterns emerged from high contaminant occurrence States and low contaminant occurrence States, particularly for the organic chemicals — VOCs and SOCs. Few consistent trends were evident for the IOCs.

The proportion of systems with detections of VOCs increases with system size, from small to large systems, particularly for ground water systems. The same trend is evident for surface water systems, but it is not as consistent for all contaminants. In a survey conducted in the 1980s, EPA also found that large ground water systems had significantly greater frequency of VOC occurrence than small systems (Westrick et al., 1984).

The trend is more evident for percent detections (>MRL) than for the percentage of systems with MCL exceedances (>MCL). This can be explained, in part, because the number of systems with exceedances becomes so low that they may only occur in one size class in a State. The URCIS data, with its greater national representation, suggest that for ground water systems the proportion of systems with VOCs >MCL also increases with system size (Fig. V.A.3; Tables V.A.6-8 -- see Appendix D).

The SDWIS MCL violation data shown in Table V.A.9 illustrate this point. Again, there is a consistent increase in the percentage of systems with VOC violations from small to large ground water systems. While the patterns are not as consistent for other contaminant groups, there are no cases where the smallest systems (<500 served) have the greatest percentage of violations. The one exception in all State data reviewed, was when a system serving <500 persons was the only size category (or at most, one other category) that detected the contaminant.

A) High-occurrence State



No systems with xylene concentrations above the MCL

B) Low-occurrence States

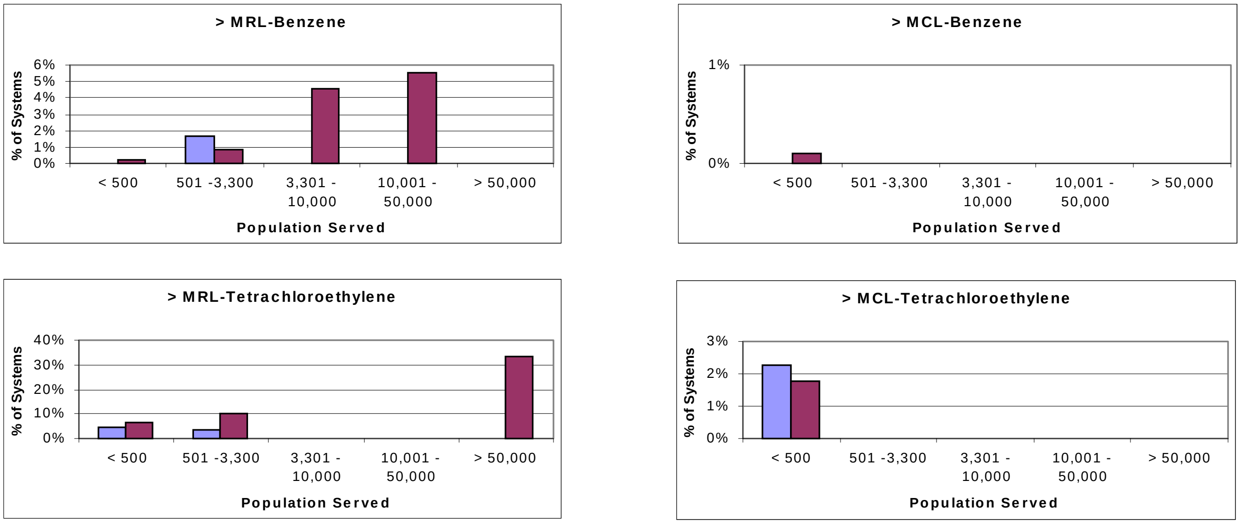


Figure V.A.1. Summary of the percentage of systems with detections (>MRL) and exceedances (>MCL) of selected VOCs, comparing ground-water and surface-water supplied systems, by size of system (population served categories), for A) a high-occurrence state, and B) a low-occurrence state.

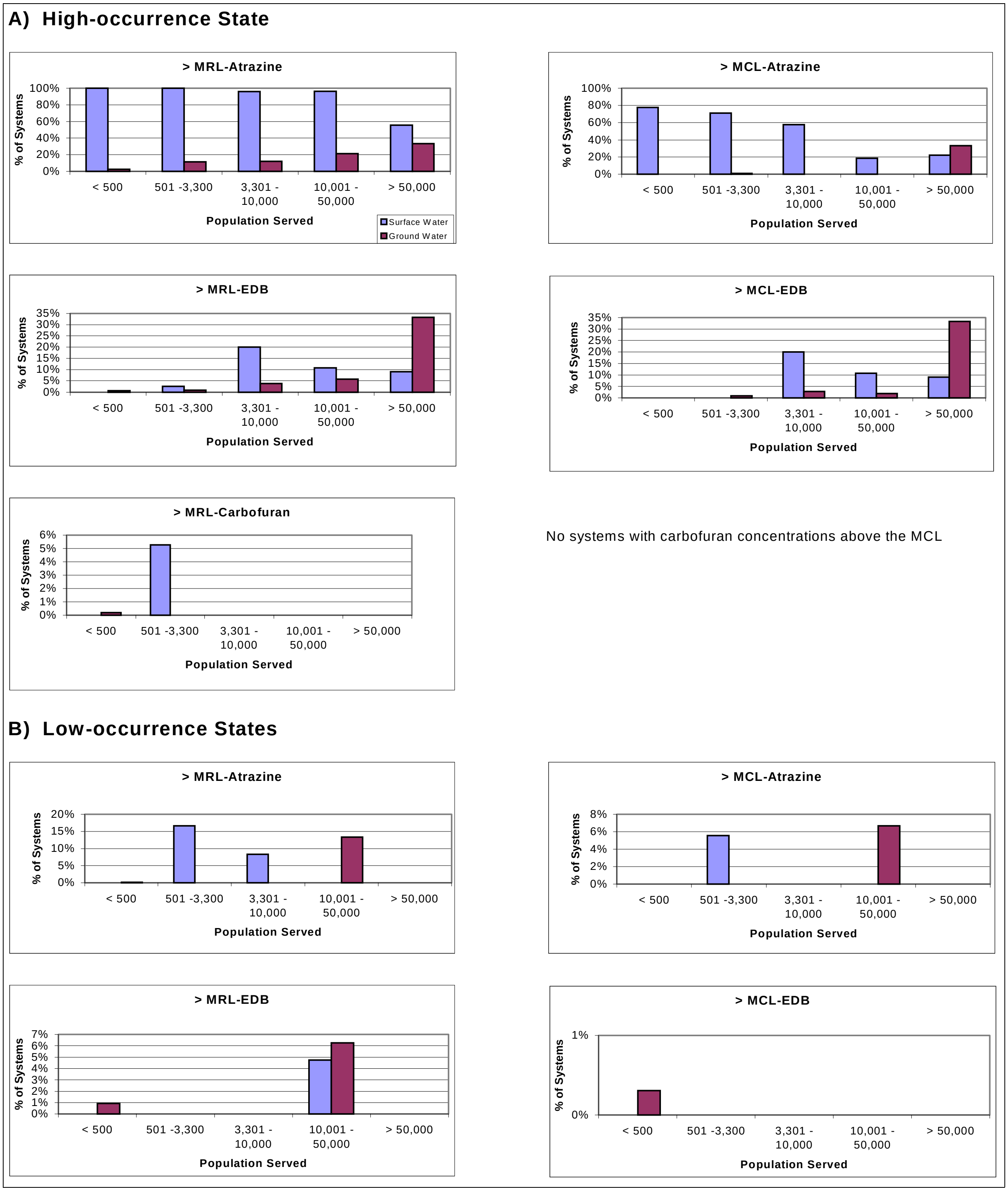
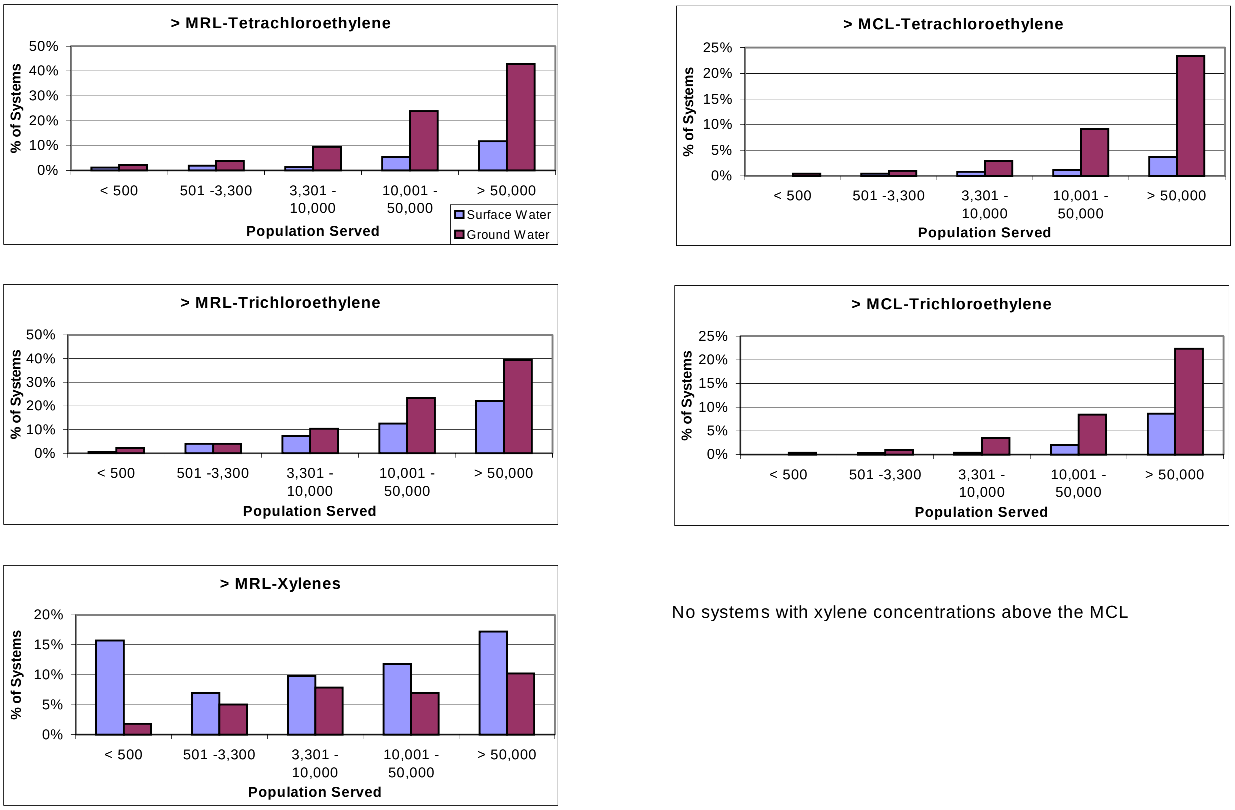


Figure V.A.2. Summary of the percentage of systems with detections (>MRL) and exceedances (>MCL) of selected SOCs, comparing ground-water and surface-water supplied systems, by size of system (population served categories), for A) a high-occurrence state, and B) a low-occurrence state.

A) VOCs in URCIS



No systems with xylene concentrations above the MCL

B) SOCs in URCIS

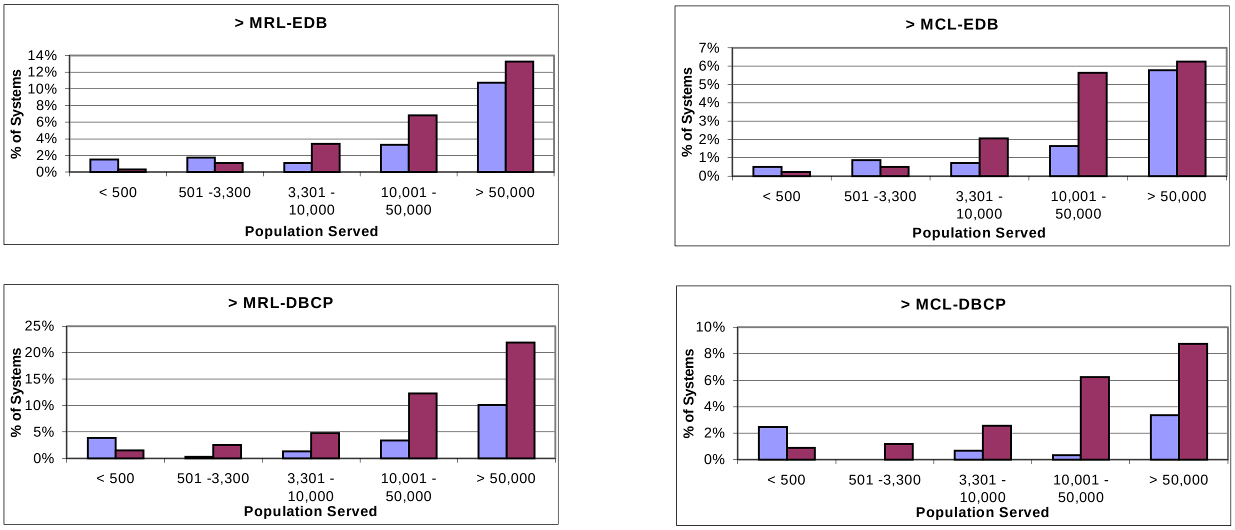


Figure V.A.3. Summary of the percentage of systems with detections (>MRL) and exceedances (>MCL) of selected A) VOCs, and B) SOCs, comparing ground-water and surface-water supplied systems, by size of system (population served categories), from the national URCIS database.

Table V.A.9. Summary of national MCL violations by system size, for surface water and ground water systems. Percent MCL violations derived from SDWIS information for 1/1/93-3/31/1998.

System Size	**Percentage of water systems nationally, with MCL violations summarized for contaminant groups, by source water and system size (population served)**					
(Population served)	**IOCs**		**SOCs**		**VOCs**	
	SW	**GW**	**SW**	**GW**	**SW**	**GW**
<500	0.2%	0.5%	0.5%	0.1%	0.2%	0.4%
501-3,300	0.4%	0.5%	1.7%	0.1%	0.2%	0.3%
3,301-10,000	0.3%	0.9%	1.2%	0.1%	0.0%	0.7%
10,000-50,000	0.5%	1.7%	0.5%	0.1%	0.4%	1.8%
>50,000	0.0%	0.0%	0.3%	0.0%	0.3%	1.1%
All	0.3%	0.6%	1.0%	0.1%	0.2%	0.4%

This same trend in ground water is apparent for many SOC contaminants, but not all (Fig. V.A.2 and V.A.3). For high-occurrence States and URCIS — both of which have more data — increasing occurrence from small to large systems is clearly illustrated. However, there are often too few systems with SOC detections for a pattern to emerge. For example, in Figure V.A.2A, carbofuran only occurs in one system size category. For a low-occurrence State (Fig. V.A.2B), EDB only occurs in two size classes of ground water systems and is predominant in systems serving >10,000 persons, but in URCIS (Fig. VA.3) the trend from small to large is clearly defined, both for occurrence (>MRL) and exceedances (>MCL).

Patterns for surface water systems are variable, sometimes increasing, sometimes decreasing with system size, again, in part because there are relatively few systems with exceedances. The pattern for systems with exceedances (>MCL) generally follow those for detections. For atrazine, on Figure V.A.2, the largest ground water systems have the greatest proportion of systems >MCL, but the trend is the opposite for surface water systems.

The data clearly indicate that, on a proportional basis, small systems do not show a greater incidence of contaminant occurrence. For many contaminants, the larger systems show greater occurrence and a tendency, at least for VOCs, to have proportionately more systems with exceedances. This trend is most consistent for ground water systems and is a logical pattern in many respects. First, systems using ground water to supply larger populations must produce greater volumes of water. To pump large, consistent volumes of ground water most systems use relatively shallow, often unconfined, aquifers that can produce the water needed. Many larger systems use alluvial aquifer well fields. As already described, shallow, unconfined aquifers, such as alluvial aquifers, are also the ground water settings most vulnerable to contamination. The greater pumping rates may further contribute to greater contributing area to these wells. Many times, larger systems serve and are located around major population centers, which typically have a greater number and density of contaminant sources, particularly for VOCs. Many monitoring requirements are more stringent for systems serving larger populations, in large part because they do serve the largest proportion of the population. The data support that general approach.

V.A.1. Systems By The Numbers; Other Perspectives On Occurrence

The percentage of systems with occurrence by size or source water categories provides insight to the relative, inherent vulnerability to contamination. There are other important perspectives in these numbers that also should be kept in mind, such as the number of systems involved or the population affected. These are briefly discussed below.

V.A.1.a. System Numbers Perspective. Table V.A.10 summarizes the number of systems within each system-size/source-water category (for CWSs and NTNCWSs). In the U.S. there are many more small systems than large systems. For example, there are approximately 46,000 ground water systems serving <500 people and about 200 serving >50,000 people. One percent of ground water systems serving <500 people represents a larger number of systems (~460) than the total number of ground water systems serving >50,000 people. Thus, lower proportional occurrence in small systems may still translate into a greater number of systems with problems, directly impacting the resources needed for program oversight for States. Tables V.A.11 and V.A.12 provide perspective on the *number* of systems with contaminant detections in relation to the percentages of systems.

Table V.A.11 illustrates the perspective by CWS and NTNCWS system size, by surface water and ground water for the occurrence of PCE (tetrachlorethylene). For this example the number of systems with detections is calculated by multiplying the percentage of systems with detections from the URCIS database times the national inventory numbers in Table V.A.10. As described previously, on a proportional basis, there is a substantial and consistent trend with occurrence increasing from small to large ground water systems: 43% of the largest ground water systems (those serving more than 50,000 people) have detections of PCE while only 2% of the smallest ground water systems (those serving less than 500 people) have detections. This indicates that 87 of the largest ground water systems detect PCE while approximately 1,000 of the smallest ground water systems detect PCE, even though there are proportionately 40% fewer systems with detections in this size class. Hence, while the proportion of systems with detections steadily increases with increasing size, the number of systems with detections declines.

The comparative perspective of percentage of systems and number of systems also is pertinent to the comparison of source-water categories. Table V.A.12, shows the summary comparison of surface water and ground water systems from the national cross-section (after Table IV.C.3), including both percentages and number of systems nationally. The number of systems is calculated by multiplying the percent occurrence from the national cross-section times the national inventory (Table V.A.10). For atrazine there is a significant difference in vulnerability: 21% of surface water systems have detections, compared to 2% of ground water systems. This translates to nearly identical number of systems—1,275 for surface water and 1,247 for ground water—because of the 10-fold difference in system-source water numbers. For xylenes, the percentages still indicate greater vulnerability for surface water systems: over 12% detecting xylenes, compared to 4% for ground water. From a program implementation standpoint, this equals about 740 surface water systems vs. 2,400 ground water systems, however.

Both the proportional occurrence and the actual number of systems with problems are important perspectives to understanding the implications of these data. The percentage of systems with occurrence of a contaminant by system size or by source water illustrates the relative vulnerability of systems. The actual number of systems affected directly relates to the level of effort required for program implementation, the necessary oversight, and the total costs involved.

V.A.1.b. Population Perspective. Exposure assessment is not the focus of this report, however, as noted in previous sections, one of the most important perspectives on occurrence and system size is related to population exposure. A relatively small portion of the approximately 68,000 CWSs and NTNCWSs serve the majority of the national population. As shown on Table V.A.10, 96% of the total CWSs and NTNCWSs are small systems, each serving less than 10,000 people, while 4% are large systems, each serving greater than 10,000 people. However, the 4% comprised of large systems serve nearly 80% of the population, while the 96% comprised of small systems, serve about 20% of the population. Hence, the increased proportion of large systems with occurrence of contaminants also translates into a disproportionately greater portion of the population being exposed. All of these perspectives on occurrence have bearing on considerations for design of monitoring schemes and other facets of drinking water programs.

Table V.A.10. Total number of nonpurchased, community and non-transient non-community water systems (CWSs and NTNCWSs), by source water and system size. From January 1997 SDWIS database.

Size Category	**# Systems**	**# Systems**	**Total # Systems**
	Surface Water	Ground Water	TOTAL
500 and under	1,900	46,200	48,100
501 to 3,300	1,820	12,306	14,126
3,301 to 10,000	1,006	2,404	3,410
Subtotal ≤10,000	*4,726*	*60,910*	*65,636*
10,001-50,000	927	1,254	2,181
over 50,001	389	204	593
Subtotal >10,000	*1,316*	*1,458*	*2,774*
TOTAL	6,042	62,368	68,410

Table V.A.11. Community and non-transient non-community water systems (CWSs and NTNCWSs) with detections of PCE >MRL and >MCL from URCIS, by system size. Number of systems affected nationally extrapolated from SDWIS inventory in Table V.A.10.

Percentage and numbers of water systems nationally, with detections >MRL and detections >MCL* for PCE, by source water and system size (population served)								
System size (Population served)	**SW > MRL**		**GW > MRL**		**SW > MCL***		**GW > MCL***	
<500	1.1%	22	2.2%	1,001	0.0%	0	0.4%	198
500-3,300	2.0%	36	3.8%	470	0.4%	8	1.0%	126
3,301-10,000	1.3%	13	9.6%	230	0.8%	8	2.9%	70
10,001-50,000	5.5%	51	23.8%	298	1.2%	11	9.2%	115
>50,000	11.8%	46	42.8%	87	3.7%	14	23.3%	48

* % > MCL indicates the proportion of systems with any analytical results exceeding the concentration value of the MCL; it does not necessarily indicate an MCL violation. An MCL violation occurs when the MCL is exceeded by the average results from four quarterly samples or confirmation samples as required by the primacy State.

Table V.A.12. Summary comparison of occurrence of selected Phase II/V contaminants in water systems using surface water vs. ground water, from national cross-section States. Shows percentage of systems and number of systems (after Table IV.C.3).

Contaminant	Percent > MRL		# > MRL		Percent > MCL*		# > MCL	
	Surface Water	Ground Water	Surface Water	Ground Water	Surface Water	Ground Water	Surface Water	Ground Water
Individual Contaminants								
IOCs								
Barium	49.1%	47.3%	2,967	29,500	0.5%	0.2%	30	125
Cadmium	5.1%	4.9%	308	3,056	0.2%	0.6%	12	374
Mercury	9.0%	4.5%	544	3,368	0.5%	0.4%	30	249
SOCs								
2,4-D	11.2%	1.2%	677	748	0.0%	0.0%	0	0
Atrazine	21.1%	2.0%	1,275	1,247	10.7%	0.1%	646	62
Benzo[a]pyrene	0.5%	0.5%	30	312	0.0%	0.1%	0	62
Carbofuran (Furadan)	0.8%	0.1%	48	62	0.0%	0.0%	0	0
Ethylene Dibromide (EDB)	4.2%	1.0%	254	624	3.7%	0.7%	224	437
Glyphosate (Roundup)	0.0%	0.1%	0	62	0.0%	0.0%	0	0
Simazine	15.9%	1.4%	961	873	1.0%	0.0%	60	0
VOCs								
Benzene	3.9%	1.2%	236	748	0.3%	0.4%	18	249
Styrene	4.1%	2.1%	248	1,310	0.0%	0.2%	0	125
Tetrachloroethylene	7.1%	4.3%	429	2,682	1.7%	1.8%	103	1,123
Vinyl Chloride	3.1%	0.5%	187	312	0.3%	0.2%	18	125
Xylenes (Total)	12.3%	3.9%	743	2,432	0.0%	0.1%	0	62
Group Summaries								
IOCs								
IOCs-Regulated	83.7%	83.5%	5,057	52,077	2.5%	4.2%	151	2,619
SOCs								
SOCs-Group 1	21.9%	2.4%	1,323	1,497	10.7%	0.1%	646	62
SOCs-Group 2	20.4%	13.4%	1,233	8,357	0.9%	1.0%	54	624
VOCs								
VOCs-Regulated	41.1%	19.9%	2,483	12,411	8.2%	6.1%	495	3,804
VOCs-Group 1	19.5%	6.6%	1,178	4,116	0.9%	0.9%	54	561
VOCs-Group 2	11.1%	6.4%	671	3,992	2.9%	3.2%	175	1,996

* % > MCL indicates the proportion of systems with any analytical results exceeding the concentration value of the MCL; it does not necessarily indicate an MCL violation. An MCL violation occurs when the MCL is exceeded by the average results from four quarterly samples or confirmation samples as required by the primacy State.

IOC-Regulated: includes all the regulated IOCs.
SOCs-Group 1: includes alachlor, atrazine, and simazine.
SOCs-Group 2: includes bis(2-ethylhexyl)phthalate, bis(2-ethylhexyl)adipate, and benzo(a)pyrene.
VOCs-Regulated: includes all the regulated VOCs
VOCs-Group 1: includes benzene, ethyl benzene, toluene, and total xylenes (LNAPLs).
VOCs-Group 2: includes cis-1,2-dichloroethylene, trans-1,2-dichloroethylene, 1,1-dichloroethene, tetrachloroethylene, trichloroethene, and vinyl chloride (DNAPLs).

V.B. Contaminant Sources and Occurrence

The prevalence of contaminants in drinking water can often be related to the use of chemical contaminants or the natural occurrence of minerals and metals within the recharge zone or watershed of a public water system. The relative contaminant source contribution should be considered when States revise their monitoring programs. Little information was collected for this study to specifically address this issue. Some observations can be summarized to illustrate the effects of contaminant sources on the occurrence of contaminants in drinking water. As noted, the occurrence of IOCs can vary greatly related to natural geologic sources and inter-related mining activities. Often such sources are localized (sub-state) and such analysis is beyond the scope of this study. At the State level patterns are not readily evident. Hence, the analysis focuses on the organic chemicals.

Table V.B.1 summarizes the average drinking water occurrence rates of selected VOCs for 8 different States, with their corresponding national ranking for total TRI releases (in pounds) per square mile. There is a consistent increase in the percentage of water systems with VOCs >MRL from low TRI States to high TRI States. The trend is best expressed by the VOC Groups 1 and 2 summaries. For individual compounds it is less consistent, but still evident. Again, the patterns are most consistent for ground water systems.

Table V.B.2 summarizes the average drinking water occurrence rates of selected SOC's for 8 different States with their corresponding ranking for total agricultural chemical expenditures. Here there is no evident consistency. While the very highest occurrence values (e.g., SOCs Group-1, or 2,4-D) are in the number two ranked State (i.e., high chemical expenditures, high pesticide use State), the highest ranked State has generally low values for occurrence. The pattern for occurrence in relation to agrichemical expenditures (and total pesticide use) differs for different compounds, largely because individual pesticide use varies greatly depending on crop type. For example, the compounds used in Illinois for corn and soybean production are not the same as those used for specialty crops and fruits in California. Illinois shows 96% of surface water systems with detections of atrazine, but California only shows 1%. Illinois ranks first nationally in atrazine use, California is not even among the top 20 States for atrazine use. In addition, many widely used pesticides are not included in SDWA monitoring requirements. The pesticide compounds in use are much more variable than the general nature of VOCs related to urban and industrial contaminant sources. Thus, it is difficult to establish clear associations between SDWA contaminant occurrence patterns and aggregate pesticide use.

Where more details are known for a particular pesticide, more useful relationships may be apparent. Figure V.B.1 shows the percentage of CWSs in a State with detections of atrazine plotted against the ranks of States by amount of atrazine applied to all crops for the twenty highest atrazine-use States (data from the Novartis study; Clarkson et al., 1997). For surface water there is a strong linear relationship. Atrazine is land applied, over large areas, which is why such a relationship is apparent. The graph also illustrates the much lower occurrence in, and lower vulnerability of, ground water systems. With the lesser occurrence values, in relation to the complexity of ground water settings, there is no significant trend for ground water systems.

The USGS NAQWA studies also have summarized some aspects of contaminant source factors related to occurrence (URL for the NAWQA National Synthesis: http://wwwrvares.er.usgs.gov/ nawqa/natsyn.html; see Appendix B). Many of the compounds included in their studies are not regulated for drinking water, but summary observations are illustrative of considerations for evaluating risks to drinking water systems. In their national synthesis to date, they have found that herbicides are more common in water sources in agricultural areas and insecticides are more common in urban watersheds. Two times as many insecticides are detected in surface water as ground water. (Insecticides tend to be less mobile and less soluble than many herbicides.) Atrazine was the most commonly detected pesticide in both surface and ground water, which coincides with the national cross-section estimates, as well. Also illustrated is that various pesticide compounds are more prevalent in shallow ground water, in different agricultural regions, related to crop type and the types of pesticides used (Kolpin et al., 1998). Petroleum-related VOCs occurred much more frequently in shallow wells in urban areas than agricultural areas (Squillace et al., 1998).

While not surprising, for the first time substantial information is becoming available to support and refine such observations. These can be useful factors to guide monitoring reform considerations.

Table V.B.1. Summary of the percentage of systems with detections (>MRL) of selected VOC contaminants, comparing surface-water and ground-water supplied systems, by the State ranking for total TRI release (in pounds per square mile); States range from low (fewest pounds released) to high (most pounds released) rank, from left to right.

Contaminant	**State Rank-Toxic Release Inventory**							
	40	**39**	**38**	**34**	**16**	**11**	**8**	**7**
Surface Water (% > MRL)								
VOCs-Groups								
Group 1	5.9%	3.8%	8.0%	7.6%	30.5%	39.3%	30.3%	42.9%
Group 2	0.0%	5.1%	6.4%	1.9%	0.0%	22.2%	30.3%	25.7%
VOCs-Individual								
Benzene	0.0%	0.6%	0.0%	0.0%	1.7%	17.1%	3.0%	2.9%
Tetrachloroethylene	0.0%	2.8%	5.0%	1.9%	0.0%	16.2%	15.2%	12.9%
Ground Water (% > MRL)								
VOCs-Groups								
Group 1	6.3%	2.2%	2.5%	2.0%	5.6%	19.5%	9.9%	28.6%
Group 2	2.4%	3.6%	5.3%	2.1%	3.4%	8.8%	13.0%	15.4%
VOCs-Individual								
Benzene	0.9%	0.5%	0.5%	0.3%	1.0%	4.2%	1.8%	2.5%
Tetrachloroethylene	2.0%	6.9%	4.0%	1.8%	1.8%	5.5%	5.7%	9.7%

Group 1 VOCs-Benzene, Ethyl benzene, Toluene, Xylenes (total)
Group 2 VOCs-cis-1,2-Dichlorethylene, trans-1,2-Dichlorethylene, 1,1-Dichloroethene, Tetrachloroethylene, Trichloroethene (Trichloroethylene), Vinyl chloride

Table V.B.2. Summary of the percentage of systems with detections (>MRL) of selected SOC contaminants, comparing surface-water and ground-water supplied systems, by the State ranking for Agricultural Chemical Expenditures; States range from low (least expenditures) to high (most expenditures) rank, from left to right.

Contaminant	**State Rank-Agricultural Chemical Expenditures**							
	40	**37**	**34**	**26**	**22**	**18**	**2**	**1**
Surface Water (% > MRL)								
SOCs-Groups								
Group 1	0.0%	0.0%	0.0%	13.2%	0.0%	7.0%	96.3%	1.3%
SOCs-Individual								
2,4-D	0.0%	0.0%	0.0%	11.8%	1.3%	0.0%	50.5%	1.9%
Carbofuran (Furadan)	0.0%	0.0%	0.0%	0.0%	1.4%	0.0%	1.8%	0.0%
Ethylene Dibromide (EDB)	6.7%	9.4%	2.0%	4.4%	0.7%	0.0%	9.4%	5.1%
Ground Water (% > MRL)								
SOCs-Groups								
Group 1	0.3%	0.0%	0.1%	3.4%	0.0%	0.5%	8.1%	2.6%
SOCs-Individual								
2,4-D	0.0%	0.0%	0.1%	4.7%	0.2%	0.0%	3.5%	1.0%
Carbofuran (Furadan)	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.1%	0.2%
Ethylene Dibromide (EDB)	2.5%	0.5%	0.0%	2.6%	0.9%	0.0%	1.3%	1.3%

Group 1 SOCs-Alachlor (Lasso), Atrazine, and Simazine

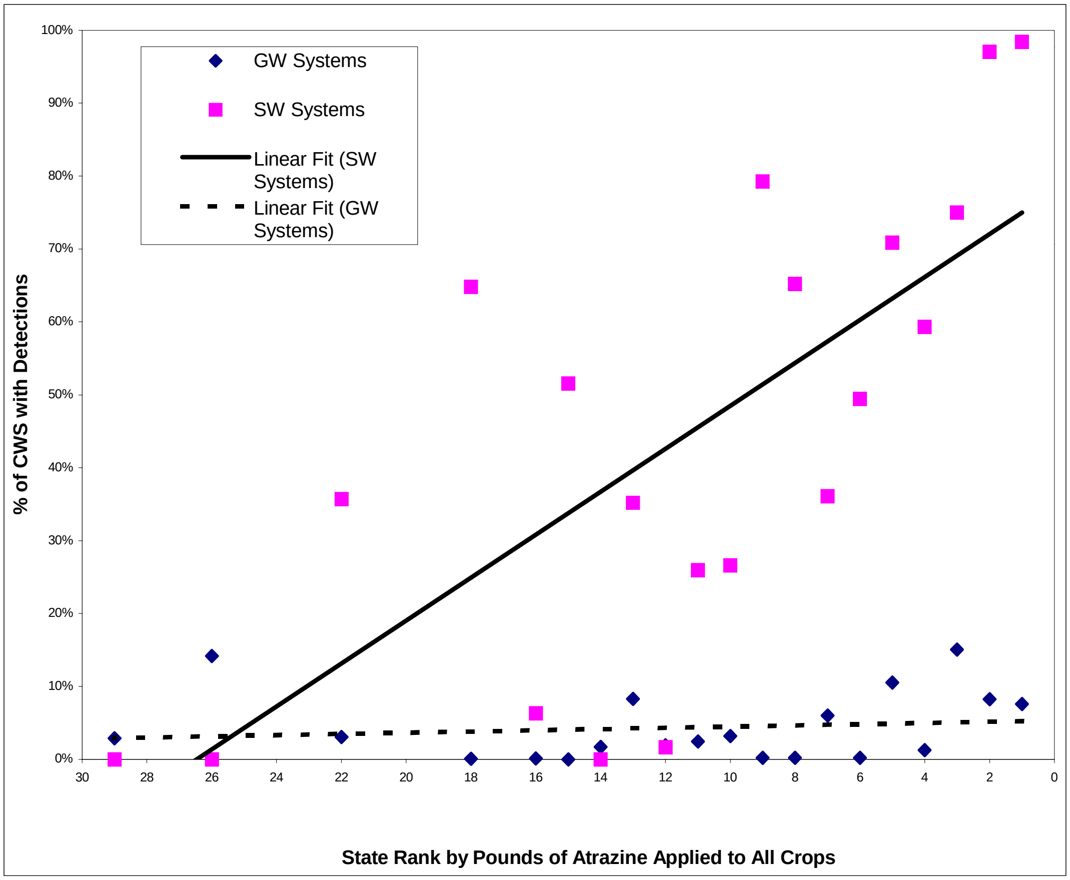


Figure V.B.1. Percentage of CWS in a State with detections (>MRL) of atrazine related to the State's national rank (1=highest rank, i.e., greatest amount applied) for pounds of atrazine applied, for surface-water and ground-water supplied systems (data from Novartis 1997).

V.C. Temporal Variability and Vulnerability

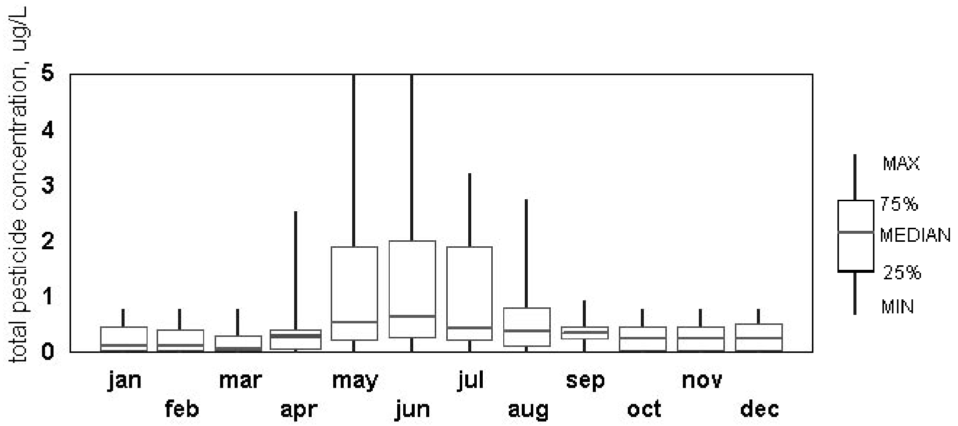
Water quality studies and monitoring throughout the United States have clearly shown that occurrence and/or concentrations for some contaminants may vary over time, both seasonally as well as from year to year. The seasonality of contaminant occurrence, or period of peak concentration, commonly varies with seasonal changes in the hydrologic cycle in relation to the source of contaminants and their fate and transport characteristics. Particularly for land-applied or land-disposed contaminants, the seasonal increase in the flux of water (e.g., spring rains) can mobilize contaminants and move them into surface or ground water flow systems. For the most vulnerable of water systems, such as surface waters, unconfined shallow ground water, and karst flow systems, for example, contaminant occurrence or peak concentrations typically occur during annual runoff and recharge periods. Targeting monitoring to these vulnerable time periods can improve the effectiveness of compliance monitoring and accuracy of exposure estimates. However, there are concerns about the cost effectiveness of seasonal targeting approaches. If, for example, many of the 64 Phase II/V contaminants exhibit different seasonal patterns, trying to seasonally target all the different contaminants could lead to a very complex and ultimately more costly monitoring regimen.

For much of the United States east of the Rocky Mountains many studies have shown the season of greatest vulnerability for contaminant occurrence is the late-spring, early-summer runoff-recharge period. This has been well established from detailed source water monitoring data, particularly for contaminants such as pesticides and nitrate (e.g., Larson et al., 1997; Barbash and Resek, 1996; Hallberg, 1989a,b). For example, Figure V.C.1 summarizes pesticide concentrations in streams from the USGS NAWQA studies. This national summary shows the concentration of pesticides in agricultural areas peaking from May through July. For streams draining urban areas the concentrations are lower and they do not show such pronounced seasonality, though May through July would still include most of the peak period.

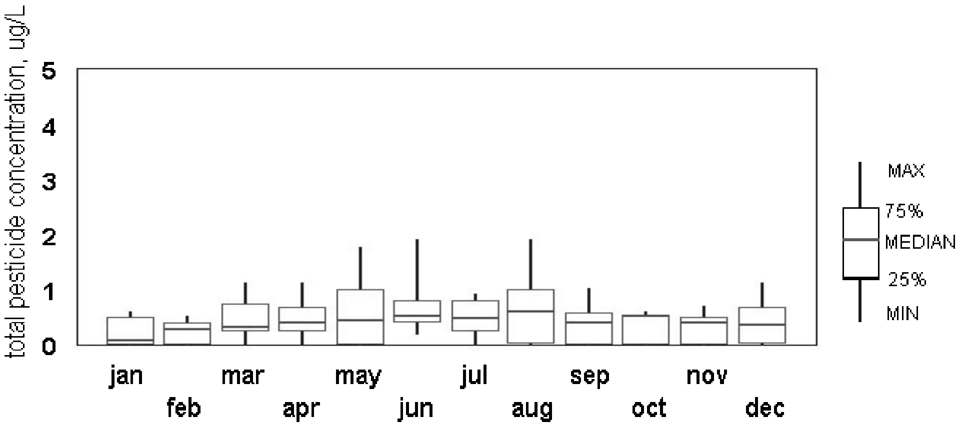
For deeper, more confined ground water systems, defining vulnerable periods is much more difficult. The exact flow path is more complex, and the time of travel much greater, and these are dependent upon many factors unique to a particular well and aquifer setting (e.g., Hallberg and Keeney, 1993). However, as depth of ground water increases (and vulnerability decreases), seasonal variability typically decreases (e.g., Barbash and Resek, 1996). There is no seasonal generality that can be applied to these deeper ground water settings.

V.C.1. SOCs

The State SDWA occurrence data were analyzed for seasonal patterns that might provide insight into drinking water monitoring schedules. Unraveling such patterns from data aggregated from many different water sources and systems is difficult, at best, as discussed below. The clearest examples are for high-occurrence pesticides. Figure V.C.2 illustrates the typical seasonal pattern for atrazine occurrence which peaks in May-July, but the number of CWSs with high monthly means decreases slowly through the fall and winter. This is one way to look at occurrence patterns.



Agricultural streams



Urban streams

Figure V.C.1. Summary of monthly average total pesticide concentrations in streams sampled in the USGS NAWQA Program, for streams affected by runoff from agricultural and urban lands. (After Larson, Gilliom, and Capel, 1999).

Figures V.C.3 and V.C.4 also illustrate seasonal patterns for pesticides, as well as the problems that can be encountered in using drinking water data to conduct such analyses. These data are from the special Ohio studies of occurrence in surface water systems (Ohio EPA, 1998). Particularly for pesticides that occur on a more intermittent basis, May-July peaks in the percentage of systems with detections are evident. For atrazine, however, the greatest percentage of systems with detections appears to be September and December. This is largely an artifact of the sampling regimen, and of the systems that were required to sample. Not all systems sampled in the fall and winter; only those systems that were known, or suspected of having year-round occurrence had to sample. Hence, in September 100% of systems sampling had detections. The seasonal occurrence pattern is more clearly defined looking at the maximum concentrations detected by month, where May-July clearly stand out.

As illustrated by this example, analyzing the State PWS data is complicated, because many sources of variation may be aggregated — many different systems, with different source characteristics, different sampling schedules, over various years, all in relation to various contaminant source characteristics. These can result in “smoothing” out the seasonal variation (e.g., percentage of systems with atrazine, Fig. V.C.3), especially for persistent contaminants that may be present all year. The aggregation of systems and source characteristics particularly confound analysis of ground water systems, but also affects analysis of surface water systems. For example, detailed studies by the USGS and others have shown that the seasonal response in reservoirs may be very different than in streams — and these are both typically identified simply as surface water sources in the PWS databases.

Small streams are more immediately affected by runoff events, and contaminant concentrations are generally greater than in large streams (which integrate a greater area). While this changes the details of temporal patterns at the daily-weekly level, the seasonal patterns are similar. Reservoirs, however, store these runoff-related events, and contaminant variations appear to be dampened. The high concentrations that enter reservoirs during runoff events may be stored for some time (e.g., months), and year-to-year variation may be more important than seasonal variations in reservoirs and lakes, depending upon reservoir size, land use in the watershed, and the turnover rate (Battaglin and Goolsby, 1998; Larson et al., 1997; Scribner et al., 1996).

Some studies have also shown secondary peak concentrations of some pesticides in fall and winter months with discharge from urban areas, but these are of much lesser magnitude than the spring period occurrence peaks (Coupe et al., 1995). Also, seasonal patterns are different in the Pacific west, for example, where fall and winter are important rainfall and recharge periods. Patterns also can be complicated by irrigation schedules or releases from irrigation storage reservoirs in the arid west (e.g., Larson et al, 1997; Kuivila and Foe, 1995).

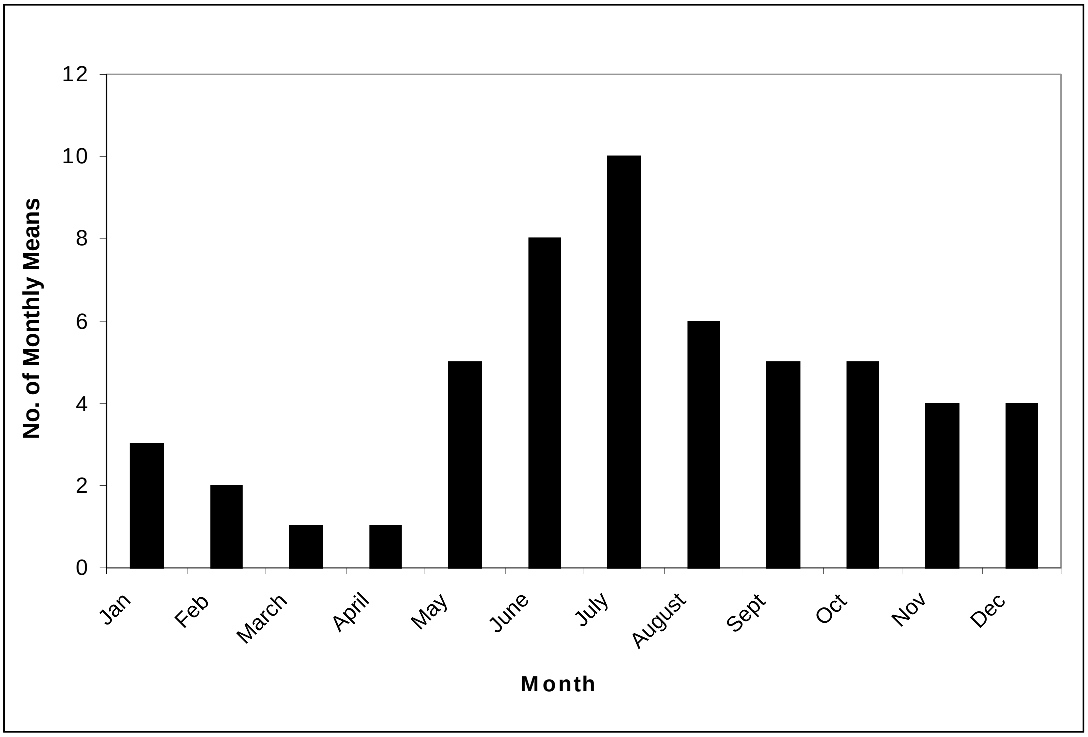


Figure V.C.2. Number of community water systems with monthly mean atrazine concentrations above 3.0 μg/L (in raw water), from special sampling study in Iowa, by Novartis Crop Production (Novartis, 1997; Clarkson et al., 1997).

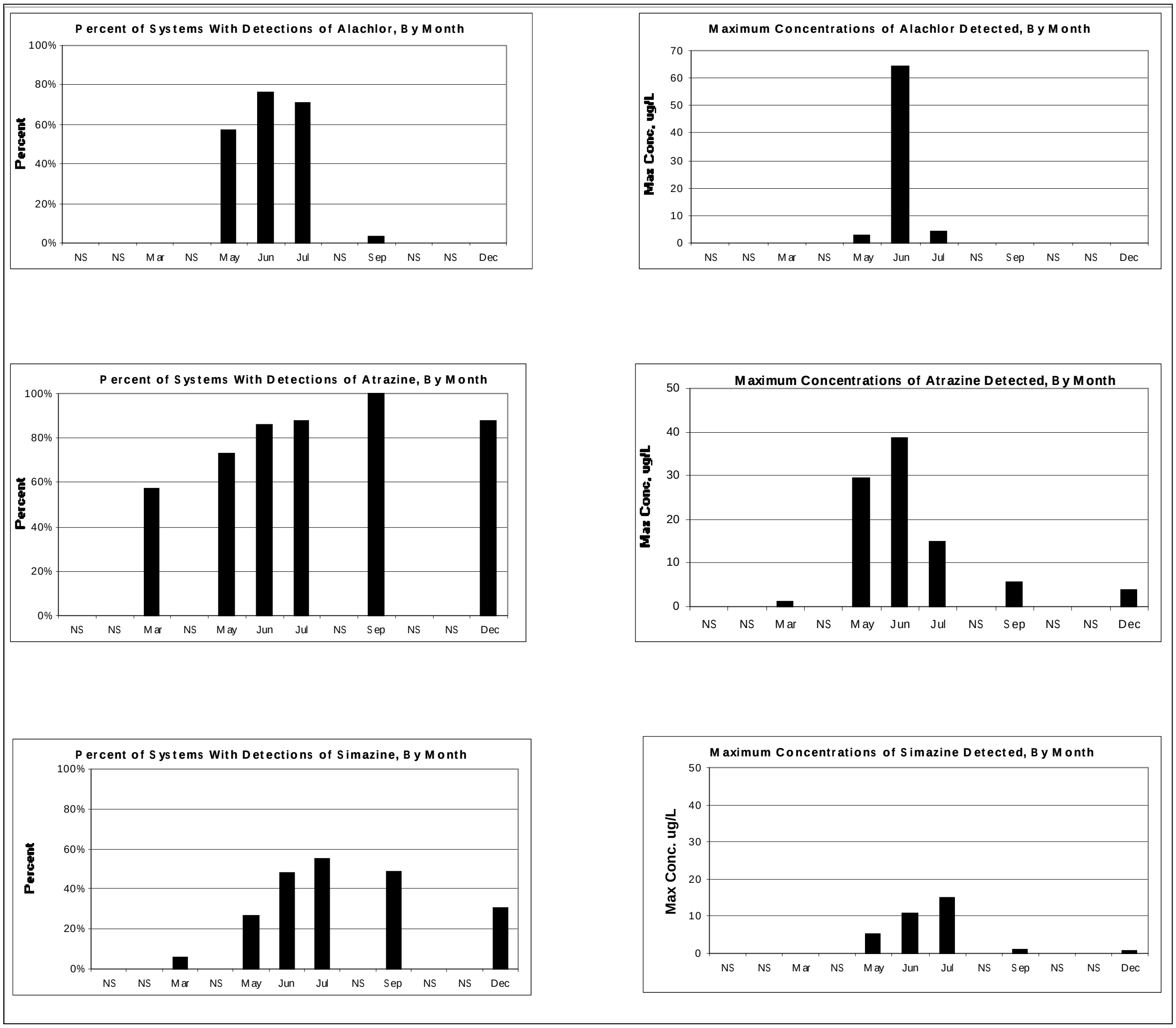
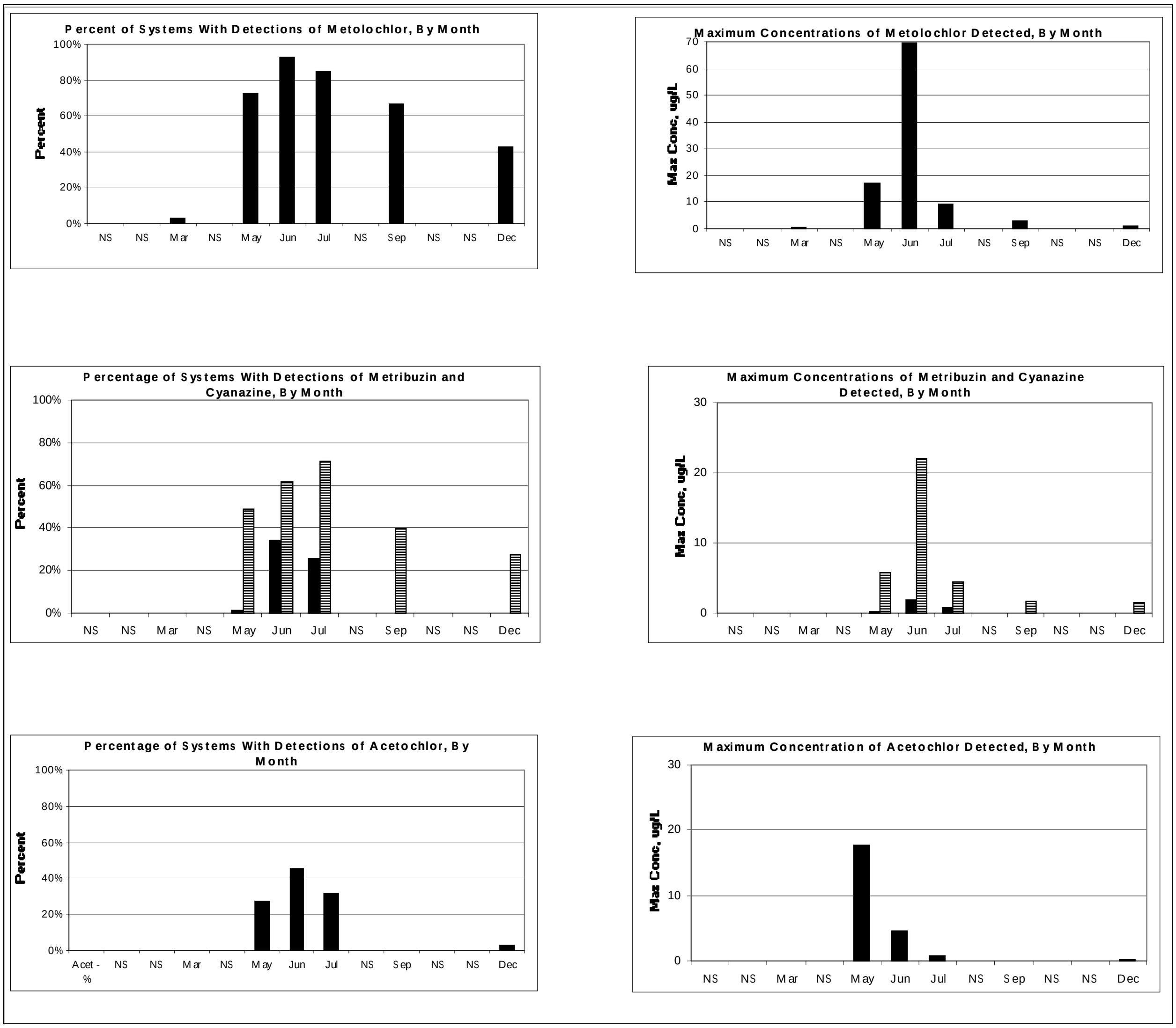


Figure V.C.3. Percentage of systems with detections and maximum concentration detected, by month, for various herbicides, for surface water systems in Ohio (Ohio EPA, 1998).



Percent of Systems With Detections of Metolochlor, By Month

Maximum Concentrations of Metolochlor Detected, By Month

Percentage of Systems With Detections of Metribuzin and Cyanazine, By Month

Maximum Concentrations of Metribuzin and Cyanazine Detected, By Month

Percentage of Systems With Detections of Acetochlor, By Month

Maximum Concentration of Acetochlor Detected, By Month

Figure V.C.4. Percentage of systems with detections and maximum concentration detected, by month, for various herbicides, for surface water systems in Ohio (Ohio EPA, 1998). (In the Metribuzin and Cyanazine chart, Cyanazine is depicted in the gray/striped column.)

V.C.2. VOCs

Many SOCs (the pesticide compounds in particular) exhibit strong seasonal patterns because their application, or discharge into the environment, is concentrated seasonally. Particularly for pesticides used in broad-scale grain production, the application season is relatively focused in the spring and early summer and coincides with annual runoff and recharge periods. This coincidence is optimal to produce seasonal patterns of occurrence in vulnerable waters. In contrast, VOCs do not typically show such seasonality in either source or discharge into the environment. Studies of individual water systems, or hydrologic settings sometimes show patterns that parallel seasonal hydrologic patterns, but on a large scale, no clear, general patterns emerge.

Figures V.C.5 through V.C.8 summarize various occurrence data by month. The VOC data, like the SOC data were analyzed several ways, including monthly number or percent of samples and systems with detections (>MRL, >0.5MCL, >MCL), the percentage of detections per month as a function of all detections or the portion of systems sampling per month; and monthly concentrations (median, 95th percentile, maximum). Even individual systems with common occurrence were isolated to assess possible temporal trends. No systematic trends were apparent. All the results look similar to the examples in Figures V.C.5 through V.C.8. There are no consistent seasonal patterns that emerge for VOCs. Figure V.C.5, shows monthly charts for xylene for several States. From the ground water systems from Illinois and the surface water systems from Michigan a 'bell-shaped' occurrence pattern, peaking in mid-summer, might be surmised. However, the Illinois surface water systems alternate peaks and declining values. Oregon shows a peak in December, although this could be a function of Oregon's different climatic regimen. Other analysis in these States suggest that the patterns are more related to what systems are sampling, rather than a seasonal pattern, especially when groups of related contaminants are viewed. For example, Figure V.C.6, shows several VOC Group-1 contaminants for one State. Hints of seasonal patterns for one contaminant are out of phase with others. For ground water systems, for xylene, an overall "bell-shaped" pattern occurs, except that the lowest month is in the middle of the bell. Figure V.C.7 and 8, show a similar lack of pattern for the heavier VOCs (PCE, and TCE), and all VOCs aggregated. The USGS has completed an in-depth review of daily VOC data from a ten-year period from water utilities along the Ohio River (Lundgren and Lopes, in press) and also have found no significant seasonal patterns for VOC occurrence. While some data suggest that light VOCs might be more prevalent in cooler months in surface waters, the patterns are not broadly evident.

While there are undoubtedly individual water systems or watersheds where seasonal patterns could be productively targeted, this would need to be developed by individual States and systems from detailed, local information. (There are likely seasonal patterns for light petroleum fuel related compounds in reservoirs or water bodies with heavy boating use, for example.) Unlike many SOCs, no general patterns for VOCs are evident on a regional, let alone a national basis.

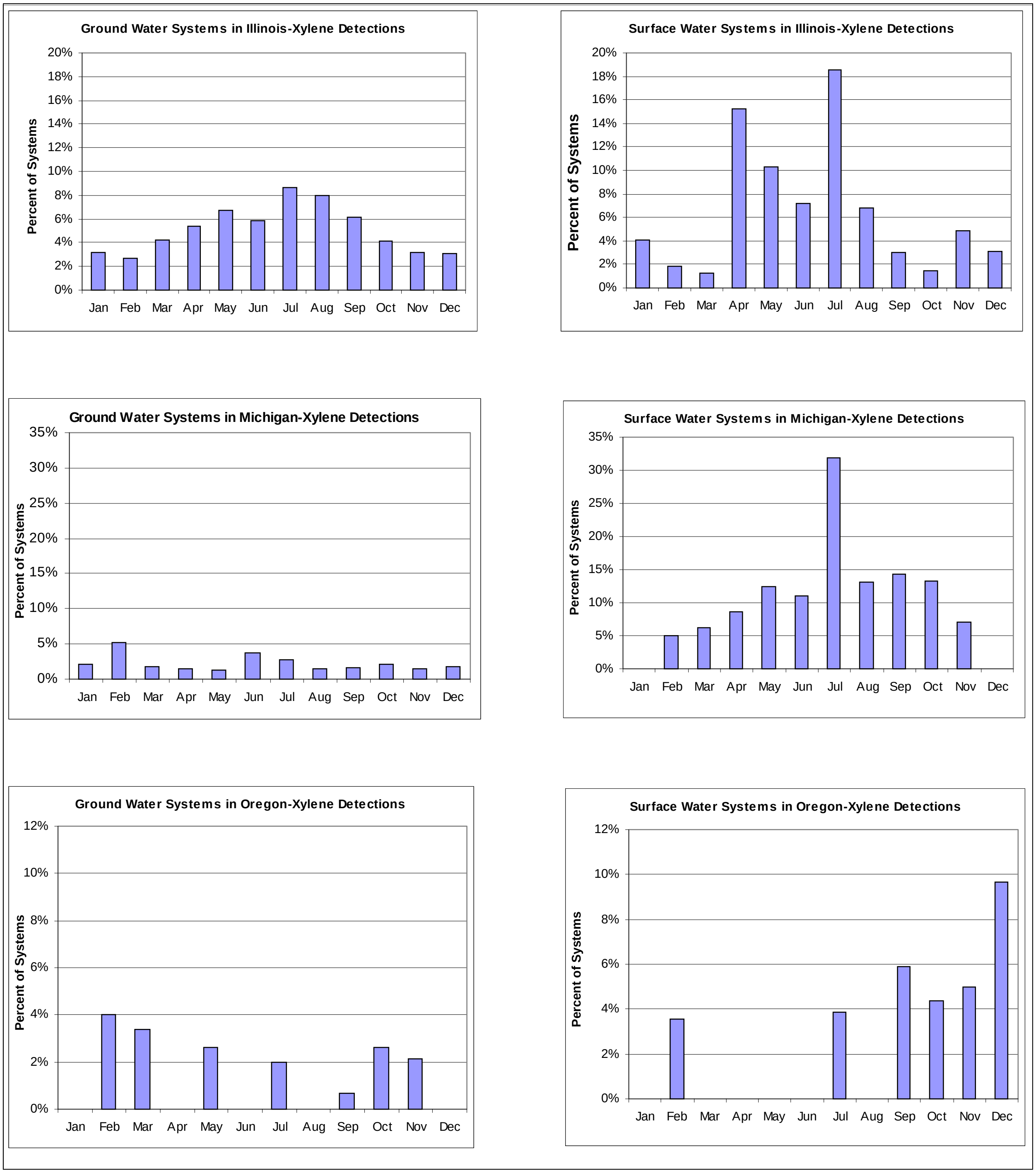
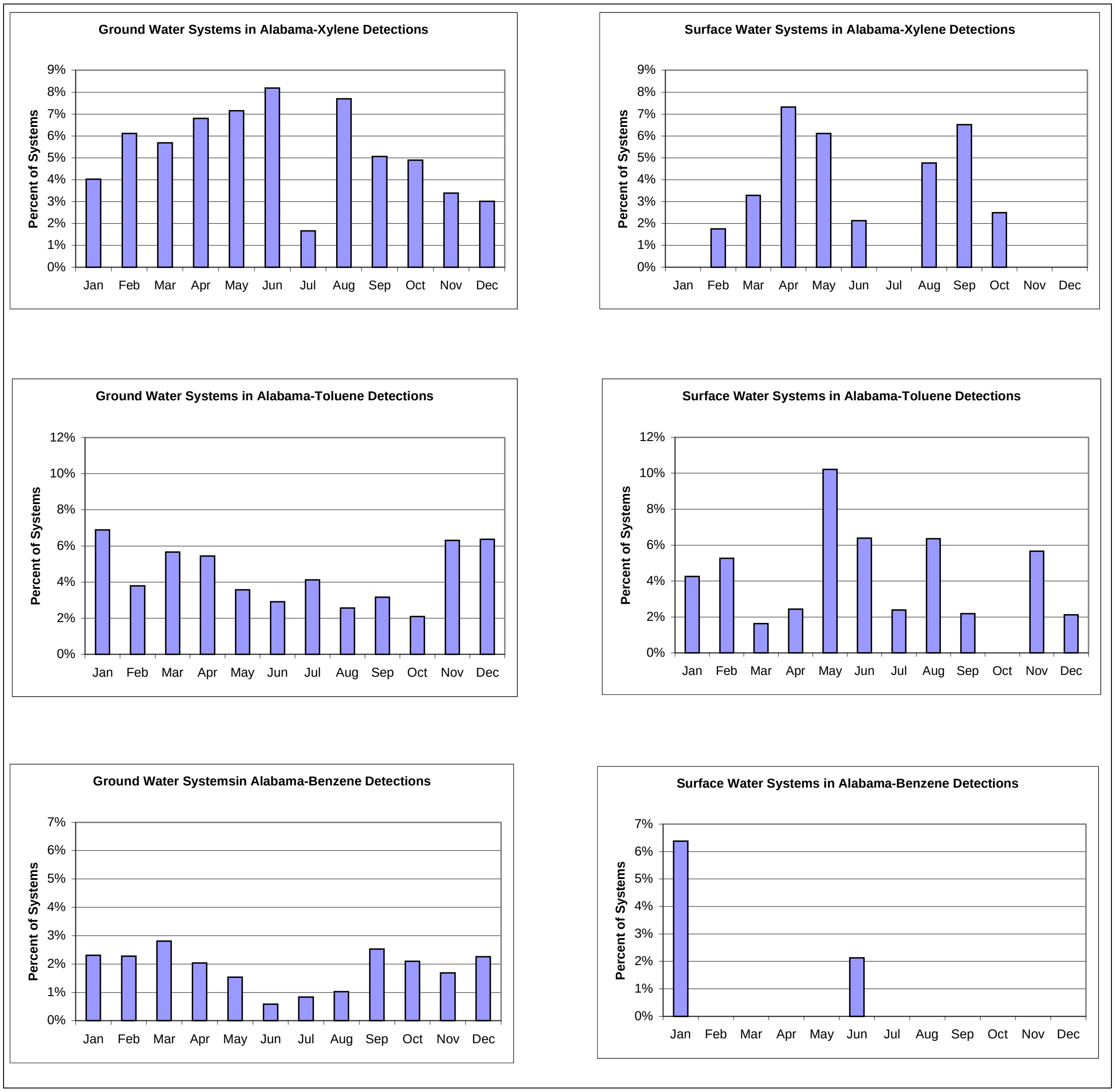


Figure V.C.5. Percentage of systems with detections of xylene, by month, for three States.



Ground Water Systems in Alabama-Xylene Detections

Surface Water Systems in Alabama-Xylene Detections

Ground Water Systems in Alabama-Toluene Detections

Surface Water Systems in Alabama-Toluene Detections

Ground Water Systemsin Alabama-Benzene Detections

Surface Water Systems in Alabama-Benzene Detections

Figure V.C.6. Percentage of systems with detections of xylene, toluene, and benzene, by month, for Alabama.

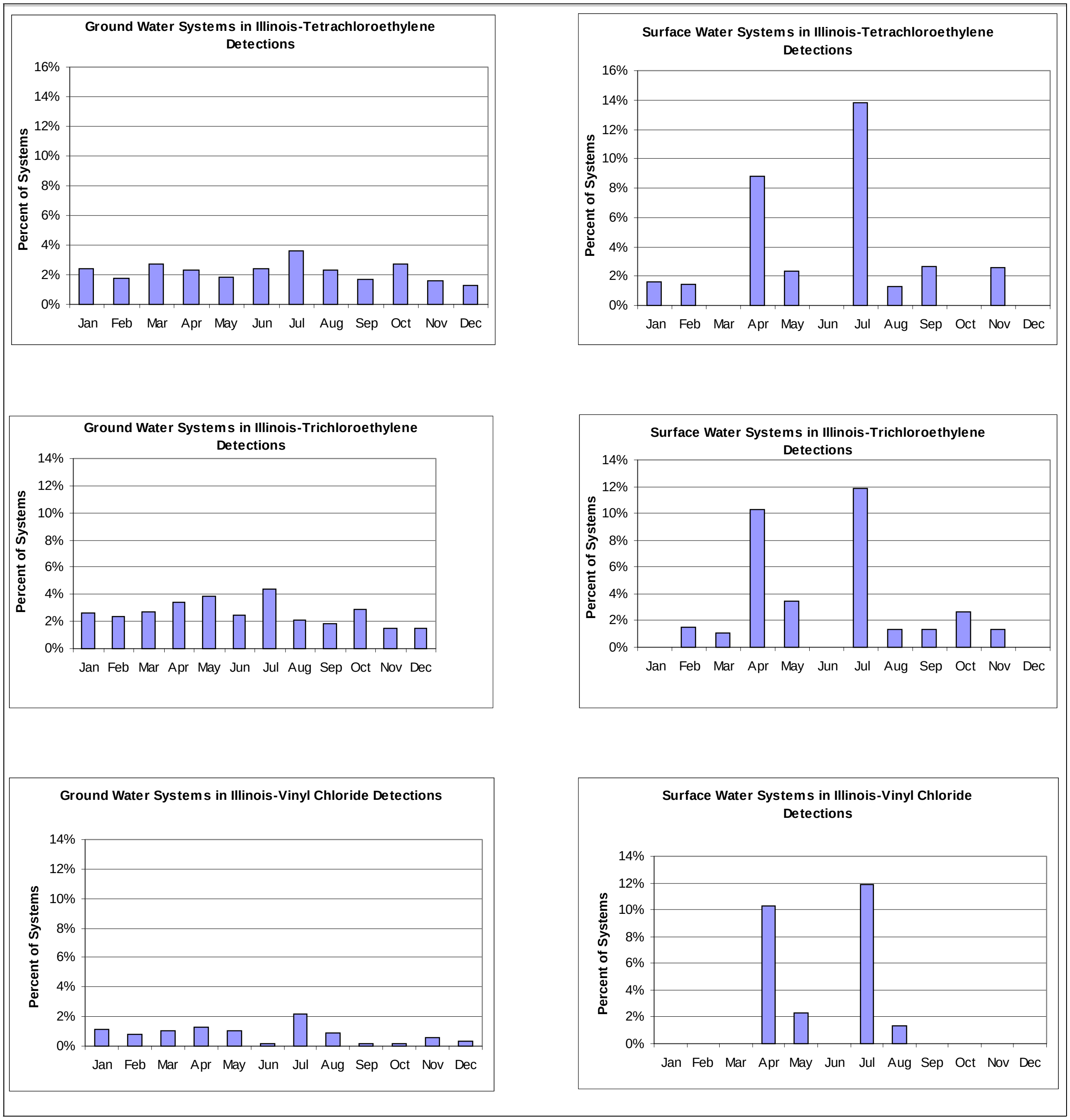


Figure V.C.7. Percentage of systems with detections of tetrachloroethylene and trichloroethylene, by month, for Illinois.

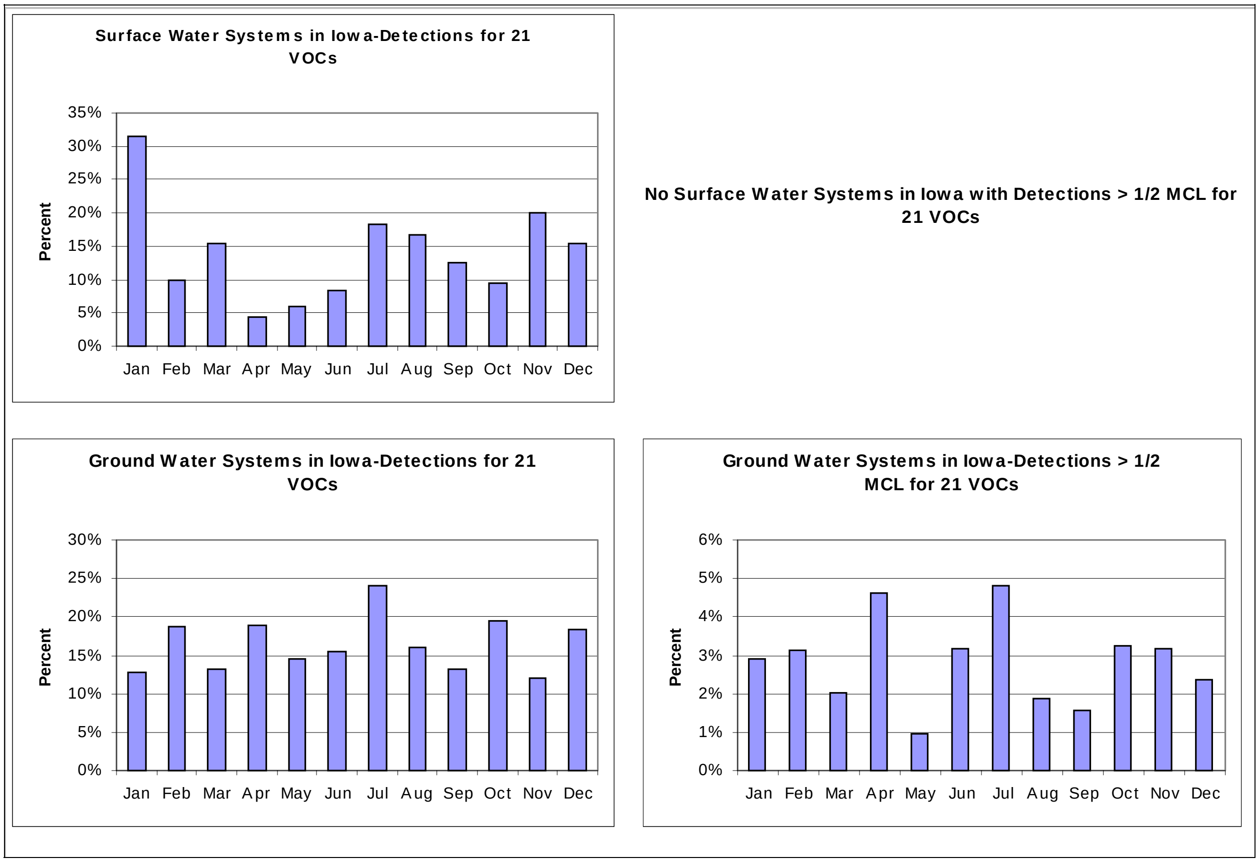


Figure V.C.8. Percentage of systems with detections (>MRL, >0.5MCL) of any of the 21 regulated VOCs, by month, for Iowa.

V.C.3. Implications For Monitoring

The impetus to target monitoring to vulnerable time periods has been the recognition that simple quarterly monitoring often does not adequately account for seasonal peak concentrations for compliance analysis or exposure estimates. Figure V.C.9 shows a schematic annual concentration record (generated from herbicide concentration data from a Midwestern river with high-ground water baseflow). This configuration is typical of the strong seasonal occurrence patterns that have been studied. Without special water treatment by the PWS, these concentration trends would also be apparent in finished drinking water that uses this source (Hallberg, 1989a; Hallberg et al., 1996). Quarterly water samples could be collected at times labeled A during the year. This sampling regimen would, by chance or by choice, significantly underestimate the annual average concentration. Even scenario B, which collects one quarterly sample during the May-July peak period, would underestimate occurrence.

Statistical studies of sampling strategies in surface water (e.g., Battaglin and Hay, 1996) show that incorporating sampling during spring and early summer runoff periods provides a more accurate representation of annual occurrence than random quarterly sampling (that can avoid these months) — as in scenario A on Fig. V.C.9. In these studies, the USGS evaluated the effects of 10 sampling strategies on the estimates of annual mean concentration of herbicides, such as atrazine. The accuracy of the strategy was computed by comparing time-weighted annual mean concentrations calculated from detailed water sampling at 17 locations with the annual mean estimated by each sampling strategy, using 1,000 Monte Carlo simulations for each strategy. (In other words, each sampling strategy was simulated using 1,000 different combinations of sampling times throughout the year.) The results were compared to a tolerance value around the actual mean from the detailed water-quality data. Pertinent results are summarized in Table V.C.1. A value of ± 0.75μg/L around the actual mean (which is 25% of the MCL of 3μg/L for atrazine) is used for the tolerance. The table summarizes the percentage of sampling simulations within the tolerance, or that are over (overestimating) or under (underestimating) the actual mean plus or minus the tolerance value. Quarterly sampling underestimated the mean in 20% of the random simulations; it was within the tolerance 63% of the time -- assuming a random distribution. (The quarterly results appear much more accurate than scenario A would imply because the random simulation results in at least one-third of the simulations collecting samples during peak months.) Monthly sampling was the most accurate, but this would increase sampling requirements three-fold from the current 4 samples/year (quarterly). However, three scenarios are nearly as accurate as monthly, and would not require additional sampling. Strategies sampling once each in May and June, and considering the other 10 months as zeros, or once in April, May, and June with 9 zeroes, or once each in April, May, June, and July with 8 zeros, range from 81% to 84% within the tolerance of the actual annual mean. A sampling scenario such as C, on Figure V.C.9, could provide a much more accurate view of drinking water quality, while still only requiring four samples per year. This scenario targets 3 samples in the typically vulnerable months, and collects a fourth sample during the off-season to provide a more complete record. With the type of concentration range often seen in strongly seasonal contaminants, the fall-winter background sample would have a similar numerical effect as the “zero” assumption in the simulations, but would provide a more continuous record.

Ground water studies (e.g., Pinsky et al., 1997) suggest that the more vulnerable ground water settings also show peaks during these periods, and such targeting could also be appropriate in these settings. From the data and literature reviewed, such a targeting strategy for SOCs would be adequate to improve the effectiveness of compliance monitoring and accuracy of exposure estimates. Most of the data suggest that most organic contaminants will vary in the same cycle — or, as with many VOCs, will show little systematic or seasonal variability. Hence, VOCs could be sampled on a similar schedule with SOCs and not lose resolution. However, this approach will always be most effective if States and systems use their local knowledge to define seasonal vulnerability patterns, and adjust such schedules. For example, in the Pacific west, some pesticides show peak concentrations in fall-winter, because of the use of pesticides on orchards during their dormant season, this is the rain/runoff season in this climate, and, in some cases, related to reservoir release schedules (e.g., Kuivila and Foe, 1995; Larson et al., 1997).

There is no simple, single guideline that fits all situations that will improve the accuracy of monitoring while balancing it against the burden of monitoring. All the factors discussed in this report should be considered to devise strategies to refine monitoring to target high risk situations and reduce monitoring burden on those systems with lower levels of risk.

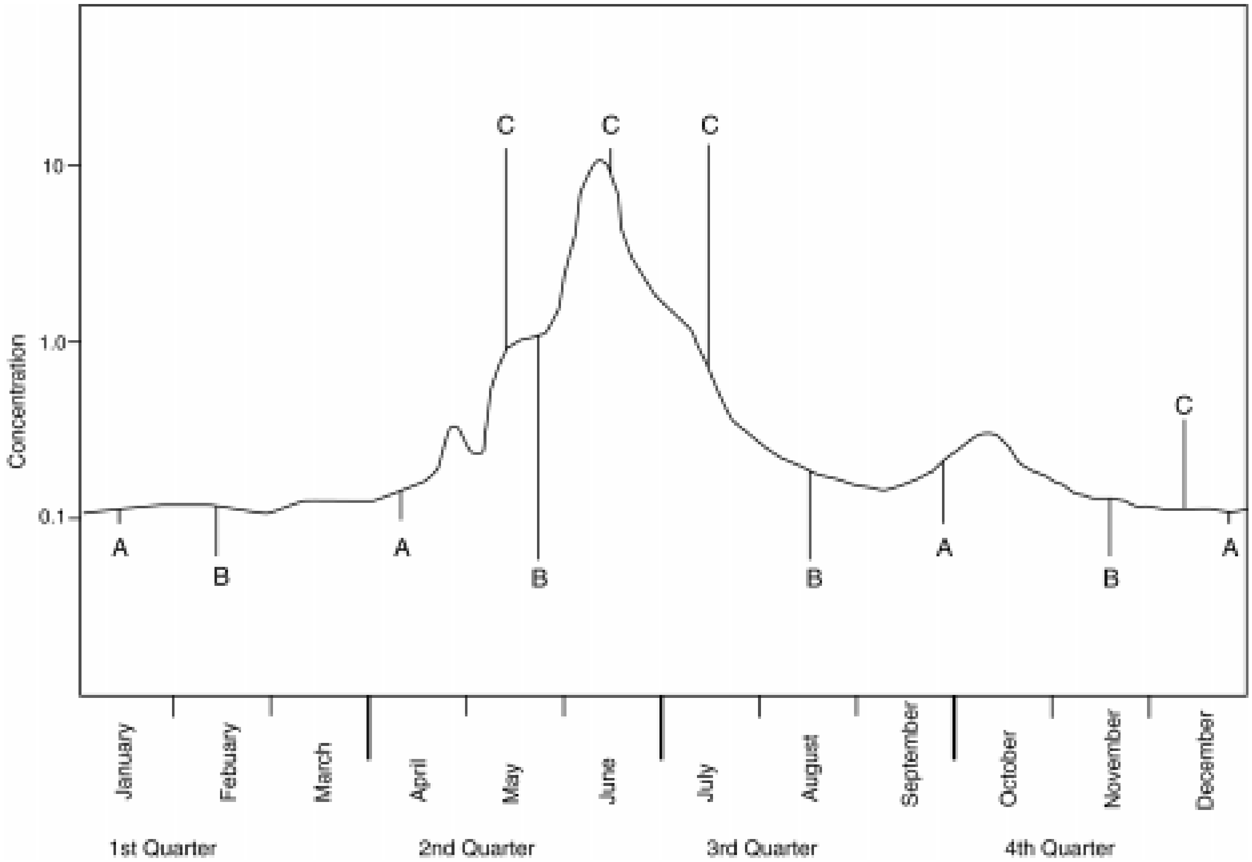


Figure V.C.9. Schematic annual contaminant concentration profile, derived from actual data from a Midwestern stream, with three sampling scenarios (A, B, and C) noted (with four sampling times for each).

Table V.C.1. Percentage of Monte Carlo sampling simulations that are within, over, or under the tolerance of the time-weighted annual mean atrazine concentration calculated from detailed field sampling. Battaglin and Hay, 1996.

Percentage of Monte Carlo simulations within, over, or under ± 0.75 μg/L of the time-weighted annual mean atrazine concentration; 1,000 simulations per sampling strategy, for all sites.

Sampling Strategy	Percentage of simulations within tolerance of ± 0.75 μg/L (± 25% of the MCL)		
	within	over	under
1 each, April, May, June, July	39%	**53%**	7%
Quarterly	63%	16%	20%
1 in June w/ 11 zeros	58%	1%	**40%**
1 each in May, June, w/ 10 zeros	81%	5%	14%
1 each in April, May, June w/ 9 zeros	82%	5%	13%
1 each in April, May, June, July w/ 8 zeros	84%	6%	9%
Monthly	**85%**	7%	8%

After Battaglin, W.A. and L.E. Hay. “Effects of Sampling Strategies on Estimates of Annual Mean Herbicide Concentrations in Midwestern Rivers.” *Environmental Science & Technology*. v. 30, pp. 889-896. 1996.

VI. CO-OCCURRENCE

An analysis of the co-occurrence of contaminants, i.e., the correlation of occurrence of pairs of contaminants, was also conducted. This was of interest to assess if various contaminants are associated with one another for consideration of monitoring scenarios, or prediction of exposure, and/or to assess if some contaminants co-occur commonly enough that one might proxy for another as a trigger for more complete monitoring. Some States, for example, have considered using the detection of certain key organic chemicals as indicators to require more complete monitoring for VOCs and SOCs.

There are two basic approaches that were considered for this assessment. The first approach would consider if two contaminants were or were not detected in the same sample (i.e., in samples collected from the same location on the same date). This is the most restrictive test that could be applied. The second approach would consider if two contaminants were or were not detected in the same system, over some period of time. This more liberal approach was taken as the first step, because occurrence of contaminants is generally quite low and co-occurrence will be difficult to evaluate and demonstrate. The less restrictive system-based analysis should yield higher statistical correlations than a sample-based analysis. This could then be used as an indicator of associations that might warrant further investigation. As discussed below, there were few significant findings from this analysis.

VI.A. Correlation Methods

An analysis of the correlation of contaminant occurrence by system was conducted for the Phase II/V contaminants and required unregulated contaminants from data from some National Cross-Section States and on the URCIS database. Conducting the analysis by system, rather than sample, removed the restriction of time from the study. Thus, an occurrence of two contaminants in the same system two years apart would produce a correlation even if the events are unrelated. However, this method will also capture associations between contaminants where one compound might be a degradation by-product of another and simply occurs in the same system at a later date.

The data were organized in such a way that if a compound was detected in a system at least once that compound was given a score of ‘one’. A contaminant which appears multiple times in a system is still counted only once. If a contaminant did not appear in a water system over the course of a sampling period that compound variable was given a score of ‘zero’. In other words, concentrations were not used or compared, just simple presence-absence. Correlation matrices (cross tabulations) were developed with all the contaminants compared to each other in paired fashion. (This produces a 64-by-64 matrix of results for Phase II/V contaminants for each State, for example.) Non-parametric correlation tests were computed (considering the simple 1-0 approach noted) using the SAS statistical program.

Spearman correlation coefficients (rho, or r_s) were computed, the values of which vary between 1 (complete concordance) and -1 (complete discordance). The correlation coefficient is an indication of how frequently two compounds occur in the same system as well as how often neither occurs in a system. When a pair of compounds occur together and don’t occur together the correlation coefficient will be high. When two compounds frequently occur (and not occur) independently of one another the coefficient will be low. If each compound only appears in a system where the other is absent, the coefficient may be negative.

VI.B. Correlation Results

Very few strong correlations were found, even though the probability values are highly significant for many associations. Every correlation coefficient greater than 0.3 was statistically significant, at $p = 0.001$. The magnitude of the p values is, in part, related to the large number of data. In essence, while there were many significant associations among contaminants the associations were not strong predictors. A correlation coefficient of less than 0.8 is not very useful as a predictor of co-occurrence. The vast majority of r values were less than 0.5 and this is analogous to an r^2 value that only explains approximately 25% of the variance.

While a high correlation coefficient is a strong indication of frequent co-occurrence within a group of systems, a low coefficient does not necessarily mean that occurrence of one compound is not related to the other. For instance, in a given State, 100% of systems which have detections of ethyl benzene also have detections of total xylenes but only 33% of systems which detected total xylenes detected ethyl benzene. The correlation coefficient in this case is only 0.57; the statistic doesn’t give any indication that, in this State, the presence of one contaminant only takes place in the presence of the other. It is also true that contaminants which are commonly found in a State's water system may be highly correlated even if occurrence is mutually unrelated.

The few stronger correlations that appeared, that were repeated in more than one State, were not particularly enlightening. There were some associations among IOCs. The IOCs occur naturally, and are more ubiquitous than most organic chemicals, so this was not surprising. The others showed associations among: SOCs-Group 1 contaminants, particularly atrazine and simazine; VOCs-Group-1, the LNAPLs or "floaters", related to gasoline compounds. Others had obvious explanations, such as associations among o-xylene and p-xylene, or o-xylene and total xylenes.

A few larger scale examples, at the contaminant group level, may be illustrative. These examples also suggest the other inferences that may be drawn from the group data. The national cross-section shows that 41% of all surface water systems have detections of VOCs (any VOC). Of these systems, 20% have detections of VOCs-Group 1, and 11% have detections for VOCs-Group 2. Group 1 and 2 total 31%. The two groups comprise the majority of the VOCs. If the total percentage of systems with detections of VOCs were 20% (i.e., less than the total of the two groups) it would suggest substantial co-occurrence between Group 1 and 2 compounds. However, the total is 41%, considerably more than the total of the two Groups, suggesting little co-occurrence.

The other relationship of interest raised by workgroup and stakeholders was the relationship of SOCs and VOCs. Table VI.B.1. summarizes some results and illustrates the variability among States, as well as the effect of low occurrence rates on assessment of co-occurrence. In the high-occurrence State, 92% of surface water systems have detections of some SOC. Hence, there has to be some substantive overlap with systems that also have detections of VOCs — particularly when 90% of systems had detections of some VOC.

Table VI.B.1. Occurrence and co-occurrence of SOCs and VOCs in two States.

% Systems with detections of:	**High-Occurrence State**		**Low-Occurrence State**	
	SW systems	**GW systems**	**SW systems**	**GW systems**
Any SOC	92%	16%	4%	3%
Any VOC	90%	66%	15%	13%
Any SOC ***AND*** VOC	82%	12%	0%	1%
Only SOCs	9%	5%	4%	2%
Only VOCs	7%	53%	15%	12%
Any SOC ***AND/OR*** VOC	99%	69%	19%	15%

Indeed, 82% of surface water systems showed co-occurrence, i.e., the systems had detections of both an SOC and a VOC. This left 9% of systems with detections of just SOCs, and 7% of systems with only VOCs. These values are additive, and in sum, 99% of surface water systems had detections of either an SOC and/or a VOC. For ground water systems there is not nearly such a strong association. Only 12% of ground water systems show co-occurrence (any SOC *and* VOC), while 58% had only an SOC or only a VOC.

Similarly, the low-occurrence State, shows the difference of occurrence patterns. For surface water systems, there is *no* co-occurrence at all between SOCs and VOCs, and only 1% co-occurrence for ground water systems. This further illustrates the generally different source relationships between SOCs and VOCs, as discussed elsewhere in this report. The analysis of co-occurrence does not provide any added insight for development of monitoring strategies. Even at the group level, the analysis indicates that SOCs and VOCs require separate consideration.

VII. SUMMARY AND CONCLUSIONS

This study provides the most extensive analysis EPA has conducted to date for the occurrence of contaminants regulated under the National Primary Drinking Water Regulations, implemented by EPA's Office of Ground Water and Drinking Water under the Safe Drinking Water Act (SDWA). The primary impetus for the study has been the interest in Chemical Monitoring Reform (CMR) in response to concerns about the efficacy of monitoring, the associated regulatory burden, and the need to consider Alternative Monitoring Guidelines (“Permanent Monitoring Relief”) for PWSs under the 1996 Amendments to SDWA. A more comprehensive review of occurrence data was needed to inform the deliberation of chemical monitoring strategy options. It is hoped that the findings and results of this study will be helpful in other assessments to make continuous improvements in drinking water programs.

Data Sources

The most important sources for this analysis were the State databases that provided SDWA compliance-monitoring data from public drinking water systems. The data includes the analytical results from Community Water Systems (CWSs) and Non-Transient Non-Community Water Systems (NTNCWSs) required to monitor for Phase II/V chemicals. Data from 14 states were reviewed. Data from 12 states were analyzed in detail. The data used in this study represent more than 10.7 million analytical results from nearly 26,000 public water systems. More than 70 percent of the data utilized date from 1993 or later; this proportion is even greater for most synthetic organic compounds (SOCs).

Four additional sources of national data were included:

- Drinking water data from systems operated by the American Water Works Service Company, which provided a multi-state perspective as a comparison to the representative cross-section results, and provided useful insight on various data quality issues.

- Regional analysis of atrazine and simazine occurrence in CWSs conducted by Novartis Crop Protection under a special study for EPA's Office of Pesticide programs, which compiled CWS monitoring data between 1993 and 1996.

- EPA's Unregulated Contaminant Monitoring Information System (URCIS) which includes data from about 40 States and territories from the first round of unregulated contaminant monitoring (1989-1993).

- USGS Studies, particularly results from National Water Quality Assessment Program. While these data were useful for reviewing national occurrence patterns, the majority of these data represent ambient water and not drinking water. These data were extracted from 75 reports and covered parts of 42 States.

Data Management and Handling

The large, unique databases used presented many challenges that could affect the resultant analysis. Hence, some data management issues warrant review. In developing an approach to data management and handling, the primary objective was to develop a consistent and repeatable analytical method that would allow valid comparisons between and among the various data sets, thereby enabling an overview of occurrence patterns at the national level. All data sets required some formatting and editing to facilitate analysis. Variables were cross-checked for consistency and to assure that each was represented or converted to represent the same units. While analysis of the data was consistent from one data set to another, each data set required some unique editing and filtering because of differences among basic data elements. These were corrected prior to analysis. All statistical analyses were conducted in SAS statistical software, with data formatting problems corrected in Microsoft Excel supplemented by specialized programs written in Visual Basic prior to analysis.

Basic data elements were reviewed to ensure consistent and comparable analyses, including:

- Contaminant Identifiers: In general, EPA method numbers were used, adjusted for unique method ID numbers used in some States.

- Detection or Reporting Limits: Minimum Reporting Levels (MRLs) for contaminant methods were also cross-checked.

- Analytical Results Elements: Special care was taken to assure that results were reported or translated to identical units.

- System and Sample Elements: These included source water type (surface water, ground water, both, purchased or other); other system features included by individual States (such as population served, system type or other locational data); and a variety of sample identifiers.

- Other Codes: These were carefully reviewed to determine their applicability for inclusion and to ensure that they were entered or converted to the same units.

In addition to the data quality and consistency issues addressed above, a number of States included unique data elements or provided a unique treatment of common elements. These were resolved in consultation with the States. Whenever errors or ambiguities could not be resolved, data were eliminated from the analysis to avoid aberrant results. However, it should be noted that very few data were excluded for these reasons, when compared to the thousands of analytical results included in these data bases.

Differences in sampling frequencies were also identified as a potential analysis problem. Systems with contaminant problems sample more frequently than those with no problems. This analysis summarizes contaminant occurrence by systems, to avoid the skew that is inherent in summaries by sample. For this analysis, only standard SDWA compliance samples were used; "special" samples, or "investigation" samples that would bias results or other samples of unknown type were not used.

Finally, there were a number of unique data quality concerns with the URCIS data. Most of the data quality problems identified would bias the data in the same, conservative direction: the occurrence of contaminants is likely overstated in both frequency and maximum concentrations. Thus, aggregate values should not underestimate occurrence. Because URCIS is such a large database, it has considerable value. The sheer quantity of data in URCIS overwhelms many minor quality problems, and the spectrum of coverage (39 States) warrants review.

Sample Representativeness

While 14 of 50 States are a substantial sample, the States with available data do not constitute a representative sample of the U.S. Therefore, various means were evaluated to construct a cross-section from the available State databases that would provide a reasonable first view of national occurrence. For this initial analysis, a simple approach was adopted that would be clear and replicable, resulting in aggregate numbers that could be easily understood. Two broad factors were considered: geographic or spatial diversity and pollution potential.

To attempt a representative view from the available State data the geographic diversity of climatic and hydrologic conditions across the United States needed to be represented. In addition, population and land use or activity factors were considered as indicators of pollution potential. Measures of manufacturing, industrial, and agricultural (pesticide use) activity were used as major indicators. From these considerations, data from 8 of the States were used to construct the national cross-section. While the national cross-section cannot be stated to be “statistically representative,” the results clearly should be indicative of the central tendency of national values for occurrence. Further, the data from the 8 Cross-Section States represent over 25% of the U.S. population using PWSs and over 20%of the PWSs sampling.

The national cross-section results were compared to the other limited national data, the URCIS results, the Novartis study, and the USGS ambient water results. In short, the aggregated State data, forming the national cross-section of contaminant occurrence compiled for this report, appears to provide a conservative but reasonable approximation of national occurrence values. The comparisons with other data suggest that the cross-section summary occurrence values are likely slightly high. Yet the data are representative enough that the summary data are valuable and some basic principles can be derived to guide further analysis and collection efforts.

Analysis Results

Data were analyzed for approximately 280 contaminants, including the 64 contaminants regulated under the Phase rules, six other regulated contaminants (e.g., nitrate, lead and copper); 34 unregulated contaminants that were required to be monitored; 14 unregulated contaminants that were considered discretionary for monitoring; and approximately 160 other unregulated contaminants. Of the 160 others, 22 are on EPA's Contaminant Candidate List, including 16 that appear on the proposed new unregulated contaminant monitoring list. The remaining 140 are other unregulated contaminants reported in the State databases. Of the 10.7 million analytical records, about 73 percent were for required monitoring.

If there is a generalization that can be made about contaminant occurrence in drinking water systems, it is that occurrence is highly variable. All 64 contaminants currently regulated under the Phase rules have been detected in drinking water systems, however, the frequency of occurrence in water systems ranges widely. Only five of the 64 contaminants, all SOCs, do not have a reported detection at greater than half their MCL. All IOCs, except asbestos, have been detected in 90-100% of systems in some States. Only three IOCs (antimony, beryllium, and cyanide) have not been detected in at least one system in at least one State (in either surface or ground water). The high occurrence rates are not surprising for IOCs because they occur naturally.

Twenty-nine of the 30 regulated SOCs have not been detected at all in some States; only phthalate occurs in every State, but its occurrence relates, in part, to sample contamination from plastics used in sampling, lab work, or system plumbing. (Because of various problems such as this phthalate, adipate, and dioxin are not included in the summary discussions.) The greatest range for any of the organic chemicals is for atrazine, which ranges from zero systems in some States to 97 percent occurrence in surface water systems in Midwestern States (up to 9 percent in ground water systems).

Fourteen of the 21 VOCs have not been detected in some of the participating States. The seven that occur in every State studied, in either surface or ground water systems, are ethylbenzene, cis-1,2-dichloroethane, tetrachloroethylene (PCE), tricloroethylene (TCE), vinyl chloride, 1,1,1-trichlorethane, and xylenes. Many VOCs occur in up to, or more than 30 percent of surface or ground water systems in various States.

Only nine of the 64 contaminants occur in less than 1 percent of either surface water or ground water systems. These are all SOCs: alachlor, carbofuran, chlordane, glyphosate, hexachlorobenzene, hexachloropentadiene, oxamyl, toxaphene, and PCBs (though PCBs have not been monitored as intensively as the other SOCs).

The analysis showed that MCL exceedances are not common. The compounds with the most common exceedances are a mix of VOCs and SOCs. Three compounds occur both on the surface water and ground water list, suggesting their widespread and common impact. The general trend that emerges is that exceedances for SOCs tend to be more common for surface water systems; for VOCs, exceedances tend to be more common for ground water systems (though the differences from surface water are not great). Fluoride makes the ground water list, but fluoride detections tend to be a result of natural occurrence and treatment additions.

Comparison of Contaminant Occurrence in Surface Water and Ground Water Systems

Fourteen percent of the regulated SOCs are detected in more than 5 percent of surface water systems in the national cross-section. In contrast, no SOCs occur in more than 5 percent of ground water systems. Nearly two-thirds of the 30 SOCs are detected in more than 1 percent of surface water systems, but over three-fourths of these SOCs are detected in only a very few (less than 1 percent) ground water systems. SOCs are far more common in surface water systems than ground water systems and surface water systems have more exceedances as well.

For VOCs, the occurrence patterns reverse to some degree, but VOCs are more common than most SOCs in both surface and ground water systems. VOCs are detected in surface water systems more frequently than is generally recognized. In addition, VOCs show much greater general occurrence in surface water systems but surface water systems and ground water systems are nearly equal for the occurrence of exceedances. Ground water systems show slightly more MCL violations for VOCs.

IOCs are about equally common (detected at concentrations greater than the MRL) in surface water systems and ground water systems, but ground water systems generally have higher average concentrations.

In short, surface water systems tend to be more vulnerable than, or equally vulnerable to ground water systems for many contaminants, but they do not necessarily have more occurrences exceeding an MCL. In particular, a greater percentage of surface water systems show detections of SOCs and surface water systems show greater occurrence of VOCs, as well.

Ground Water Vulnerability

If variability is a confounding factor for surface water systems, it is the rule for ground water systems. Ground water systems are highly variable in nature because ground water adds other dimensions to the vulnerability equation. Aquifer type and well depth are key hydrogeologic factors to ground water vulnerability.

The relative vulnerability of public water systems using ground water can be evaluated to better target and customize monitoring requirements. Because of the variability in ground water settings and system level variations, developing improved monitoring requires refined information at the sub-State and system levels. Information on aquifer characteristics, well depth, well casing and construction details, wellhead protection information, as well as system and well maintenance history are some of the factors that should be considered.

System Size and Other Vulnerability Factors

In general, the same consistent patterns emerge from States with high contaminant occurrence and from States with low contaminant occurrence. Consistent patterns were not observed for IOCs, but were for organic chemicals - VOCs and SOCs.

The proportion of systems with detections of VOCs increases with system size from small to large systems, particularly for ground water systems. While the same trend is evident for surface water systems, it is not consistent for all contaminants. The trend is more evident for percent detections (MRL) than for the percentage of systems with MCL exceedances (MCL). This can be expected, in part, because the number of systems with exceedances becomes so low that they may only occur in one size-class in a State.

This same trend in ground water is apparent for many, but not all SOC contaminants. For high-occurrence States or URCIS, both of which have more data, increasing occurrence from small to large systems is clearly illustrated. Patterns for surface water systems are variable - sometimes increasing, sometimes decreasing with system size, again, in part because there are relatively few systems with exceedances.

The data clearly indicate that, on a proportional basis, small systems do not show a greater incidence of contaminant occurrence. For many contaminants, the larger systems show greater occurrence and a tendency, at least for VOCs, to have proportionately more systems with exceedances. This trend is most consistent for ground water systems and is a logical pattern in many respects. Large PWSs need larger volumes of water. To get the larger ground water yields needed, larger systems typically use unconfined ground water supplies, that are typically shallow and more vulnerable to contaminant occurrence.

The prevalence of contaminants in drinking water can often be related to the use of chemical contaminants within the recharge zone or watershed of a public water system. The relative contaminant source contribution should be considered when States revise their monitoring programs.

Temporal Variability and Vulnerability

Water quality studies and monitoring throughout the US have clearly shown that occurrence and/or concentration for some contaminants may vary over time, both seasonally and from year to year. Targeting monitoring to vulnerable times can improve the effectiveness of compliance monitoring and the accuracy of exposure estimates. However, there are concerns associated with the cost-effectiveness of seasonal targeting patterns.

Many SOCs (the pesticide compounds, in particular) exhibit strong seasonal patterns because their application or discharge into the environment is concentrated seasonally. In contrast, VOCs do not typically show such seasonality in either source or discharge into the environment. Studies of individual water systems or hydrologic settings sometimes show VOC patterns that parallel seasonal hydrologic patterns; however, on a regional scale, no clear, general patterns emerge. While there are undoubtedly individual water systems or watersheds where seasonal patterns of VOCs could be productively targeted, this would need to be developed by individual States and systems from detailed, local information.

For SOCs, statistical and modeling data suggest that sampling strategies could be designed to better target and account for vulnerable, seasonal peak occurrence, without necessarily increasing sampling frequency and burden. Few data suggest that multiple vulnerable periods would need to be targeted in the same region. Vulnerable periods are apparent for surface waters and shallow, vulnerable ground water systems. For deeper ground water systems few generalizations can be made at the scale of this study, except to note that as well depth increases the degree of temporal variability usually decreases.

Conclusions

The data compiled and analyzed for this study comprise the largest analysis of contaminant occurrence data from Public Water Systems that EPA has conducted to date. The national cross-section developed from State drinking water databases provide a reasonable overview of the occurrence of the Phase II/V contaminants. These summary data provide an improved scientific basis for evaluating monitoring and sampling programs.

REFERENCES CITED

Barbash, J.E., and E.A. Resek. Pesticides in Ground Water, volume two of the series Pesticides in the Hydrologic System. Ann Arbor Press, Inc., Chelsea, Michigan. 1996.

Battaglin, W., and L. Hay. "Effects of Sampling Strategies on Estimates of Annual Mean Herbicide Concentrations in Midwestern Rivers." *Environmental Science & Technology*. 30: 889-896. 1996.

Battaglin, W., and D. Goolsby. "Regression Models of Herbicide Concentrations in Outflow from Reserviors in the Midwestern USA, 1992-1993." *Journal of the American Water Works Association*. 1369: 34-6. 1998.

Bloemen, H.J.Th., and J. Burn, eds. Chemistry and Analysis of Volatile Organic Compounds in the Environment. London: Chapman and Hall. 1993.

California Department of Health Services (DHS). *Drinking Water Quality Monitoring Data 1984-1996; Annual Status Report*. Monitoring and Evaluation Unit, California Department of Health Services. November 1997.

Clarkson, J.R., N.A. Hines, D.P. Tierney, and B.R. Christensen. *Human Exposure to Atrazine and Simazine via Ground and Surface Drinking Water*. Novartis Crop Protection, Inc. Novartis Study No. 696-95. EPA MRID No. 44152122. 1997.

Coupe, R.H., D.A. Goolsby, J.L. Iverson, D.J. Markovchik, and S.D. Zaugg. U.S. Department of the Interior. *Pesticide, Nutrient, Water-Discharge, and Physical-Property Data for the Mississippi River and Some of Its Tributaries, April 1991-September 1992*. U.S. Geological Survey Open File Report 93-657. 1995.

Hallberg, G.R., and D. Keeney. "Nitrate." In W.A. Alley, Regional Ground-Water Quality; Chapter 2, p. 297-322. Van Nostrand Reinhold, New York, NY. 1993.

Hallberg, G.R., D.G. Riley, J.R. Kantamneni, P.J. Weyer, and R.D. Kelley. *Assessment of Iowa Safe Drinking Water Act Monitoring Data: 1987-1995*. University Hygienic Laboratory Research Report 97-1, p. 132. University of Iowa, Iowa City, IA. 1996.

Hallberg, G.R. 1989a. "Pesticide Pollution of Groundwater in the Humid United States." In H. Bouwer and R.S. Bowman, eds., *Effect of Agriculture on Groundwater; Agriculture, Ecosystems, and Environment*. Special edition, v. 26, p. 299-367. 1989.

Hallberg, G.R. 1989b. "Nitrate in Groundwater in the United States." In R.F. Follett, ed., Nitrogen Management and Groundwater Protection; Chapter 3, p. 35-74. Elsevier Scientific Publications, Amsterdam, The Netherlands. 1989.

Hallberg, G.R., K. Woida, R.D. Libra, K.D. Rex, K.D. Sesker, B.C. Kross, L.S. Seigley, B.K. Nations, D.J. Quade, D.R. Bruner, H.F. Nicholson, J.K. Johnson, and K.L. Cherryholmes. *The Iowa State-Wide Rural Well Water Survey: Site and Well Characteristics and Water Quality*. Iowa Department of Natural Resources, Geologic Survey Bureau, Technical Information Series 23. 43p. 1992.

Howard, P.H., ed. *Handbook of Environmental Fate and Exposure Data For Organic Chemicals*. Vol. I-V. Boca Raton: CRC Press. 1989-97.

Kalkhoff, S.J., D.W. Kolpin, E.M. Thurman, I. Ferrer, and D. Barcelo. "Degredation of Chloroacetanilide Herbicides: The Prevalence of Sulfonic and Oxanilic Acid Metabolites in Iowa Groundwaters and Surface Waters." *Environmental Science & Technology*. 32(11): 1738-1740. 1998.

Kolpin, D.W., D. Sneck-Fahrer, G.R. Hallberg, and R.D. Libra. "Temporal Trends of Selected Agricultural Chemicals in Iowa's Groundwater, 1982-95: Are things getting better?" *Journal of Environmental Quality*. 26: 1007-1017. 1997.

Kolpin, D.W., J.E. Barbash, and R.J. Gilliom. "Occurrence of Pesticides in Shallow Groundwater of the United States: Initial Results from the National Water-Quality Assessment Program." *Environmental Science & Technology*. 32(5): 558-566. 1998.

Kolpin, D.W., S.J. Kalkhoff, D.A. Goolsby, D.A. Sneck-Fahrer, and E.M. Thurman. "Occurrence of Selected Herbicides and Herbicide Degradation Products in Iowa's Ground Water, 1995." *Ground Water*. 35-4, p. 679-687. 1997.

Kolpin, D.W., E.M. Thurman, and D.A. Goolsby. "Occurrence of Selected Pesticides and Their Metabolites in Near-Surface Aquifers of the Midwestern United States." *Environmental Science & Technology*. 30(1): 335-340. 1996.

Kross, B.C., G.R. Hallberg, D.R. Bruner, R.D. Libra, K.D. Rex, L.M.B. Weih, M.E. Vermace, L.F. Burmeister, N.H. Hall, K.L. Cherryholmes, J.K. Johnson, M.I. Selim, B.K. Nations, L.S. Seigley, D.J. Quade, A.G. Dudler, K.D. Sesker, M.A. Culp, C.F. Lynch, H.F. Nicholson, and J.P. Hughes. *The Iowa State-Wide Rural Well-Water Survey, Water-Quality Data: Initial Analysis*. Iowa Department of Natural Resources, Geologic Survey Bureau, Technical Information Series 19. 142 p. 1990.

Kross, B.C., G.R. Hallberg, D.R. Bruner, K. Cherryholmes, and J.K. Johnson. "The nitrate contamination of private well water in Iowa." *American Journal of Public Health*. 83, p. 270-272. 1993.

Kuivila, K.M., and C.G. Foe. "Concentrations, Transport, and Biological Effects of Dormant Spray Pesticides in the San Francisco Estuary California." *Environmental Toxicology and Chemistry*. 14 (7), p. 1141-1150. 1995

Larson, S.J., R.J. Gilliom, and P.D. Capel. U.S. Department of the Interior. *Pesticides in Streams of the United States -- Initial Results from the National Water Quality Assessment Program*. U.S. Geological Survey. Water-Resources Investigations Report 98-4222, 92 p. 1999.

Larson, S.J., P.D. Capel, and M.S. Majewski. Pesticides in Surface Waters, volume three of the series Pesticides in the Hydrologic System. Ann Arbor Press, Inc., Chelsea, Michigan. 1997.

Lide, D.R, ed. CRC Handbook of Chemistry and Physics. 77th Edition. Boca Raton: CRC Press. 1996.

Lundgren, R.F., and T.J. Lopes. *Occurrence, Distribution, and Trends of Volatile Organic Compounds (VOCs) in the Ohio River and Its Major Tributaries, 1987-96.* U.S. Geological Survey Water Resources Investigations Report 99-XXX. 1999, in press.

Majewski, M.S., and P.D. Capel. Pesticides in the Atmosphere, volume one of the series Pesticides in the Hydrologic System. Ann Arbor Press, Inc., Chelsea, Michigan. 1995.

Novartis Crop Protection, Inc. *Voluntary Atrazine Monitoring Program at Selected Community Water Systems: Iowa 1996*. Technical Report 6-97, Environmental and Public Affairs Department. 1997.

Ohio Environmental Protection Agency. *Pesticide Special Study*. Division of Drinking and Ground Waters. Columbus, 1998. Available on internet at http://www.epa.state.oh.us/ddagw/pestspst.html. 1998.

Parsons, F., P.R. Wood, and J. DeMarco. "Transformations of Tetrachloroethene and Trichloroethene in Microcosms and Groundwater." *Journal of the American Water Works Association*. 76: 56-9. 1984.

Pinsky, P., M. Lorber, K. Johnson, B. Kross, L. Burmeister, A. Wilkins, and G.R. Hallberg. "A Study of the Temporal Variability of Atrazine in Private Well Water." *Environmental Monitoring and Assessment*. 47, p. 197-221. 1997.

Ryker, S.J. and A.K. Williamson. *Pesticides in Public Supply Wells of Washington State.* U.S. Geological Survey Fact Sheet 122-96. May 1996.

Schaap, B.D., and S.M. Linhart. *Quality of Ground Water Used for Selected Municipal Water Supplies in Iowa, 1982-96 Water Years*. U.S. Geological Survey Open-File Report 98-3. May 1998.

Scribner, E.A., D.A. Goolsby, E.M. Thurman, M.T. Meyer, and W.A. Battaglin. *Concentrations of Selected Herbicides, Herbicide Metabolites, and Nutrients in Outflow from Selected Midwestern Reservoirs, April 1992 Through September 1993.* U.S. Geological Survey Open File Report 96-363. 1996.

Squillance, P.J., D.A. Pope, and C.V. Price. *Occurrence of the Gasoline Additive MTBE in Shallow Ground Water in Urban and Agricultural Areas*. U.S. Geological Survey Fact Sheet 114-95. 1996.

Squillance, P.J., M.J. Moran, W.W. Lapham, C.V. Price, R.M. Clawges, and J.S. Zogorski. "Volatile Organic Compounds in Untreated Groundwater of the United States, 1985-1995." *Environmental Science & Technology*. 33(23): 4176-4187. 1999.

Vowinkel, E.F., R.M. Clawges, D.E. Buxton, D.A. Stedfast, and J.B. Louis. *Vulnerability of Public Drinking Water Supplies in New Jersey to Pesticides*. U.S. Geological Survey Fact Sheet 165-96. 1996.

Westrick, J.J., J.W. Mello, and R.F. Thomas. "The Groundwater Supply Survey." *Jour. American Water Works Assoc*. 76, 52-59, 1984.