Evaluating Changes in Matrix-Based, Recovery-Adjusted Concentrations in Paired Data for Pesticides in Groundwater

Tammy M. Zimmerman* and Kevin J. Breen

Numerous groundwater studies in the United States and other countries have been conducted since the 1980s and have documented pesticide occurrence. According to Barbash and Resek (1996), many previous studies have not been designed to assess temporal trends in pesticide concentrations and detection frequency; hence, details are often lacking on the analytical method, detection limits, and analytic recovery. Furthermore, previous studies of pesticides in groundwater designed to evaluate differences in paired data often use tests such as the traditional sign test and the Wilcoxon signed-rank test that do not handle data with tied pairs (such as censored data where both samples in the pair are nondetections, referred to hereafter as “zero-difference ties”).

Pesticide compounds that are commonly detected in groundwater include the herbicides atrazine, metolachlor, prometon, simazine, and the atrazine degradate compound deethylatrazine (DEA) (Spalding et al., 2003; Gilliom et al., 2006; Bexfield, 2008; Postigo et al., 2010). Pesticides detected in rivers draining agricultural areas (Liu et al., 2002; McConnell et al., 2006; Bexfield, 2008; Postigo et al., 2010) are often the same as those detected in groundwater. The prevalence of these pesticides in the environment reflects a combination of climatic, agronomic, pedologic, and geologic factors that determine the rate, period, and spatial extent of application, the inherent susceptibility of the aquifer, and the persistence of compounds once they are in the hydrologic system (Lindsey and Bickford, 1999; Barbash and Resek, 1996; Johnson et al., 2001; Gilliom et al., 2006; Hancock et al., 2008; G.C. Johnson et al., 2011).

The leaching potential of a pesticide compound can be evaluated by assessing the mobility of the pesticide through the soil and the persistence of the pesticide. The pesticides with the greatest potential for transport are those that are highly soluble in water (i.e., they do not readily adhere to soil particles) as indicated by low soil adsorption coefficients ($K_{oc}$) and that are relatively persistent in the environment, as indicated by low rates of decay (long soil half-lives) (Lindsey and Bickford, 1999; G.C. Johnson et al., 2011).
Carbonate aquifers overlain by predominantly agricultural land use often are particularly vulnerable to anthropogenic contamination because of the inherent susceptibility of the bedrock and contaminant availability (Kacaroglu, 1999; Baran et al., 2008; Debrewer et al., 2008; Dalton and Frick, 2008; Coxon, 2011). Pesticide occurrence is not unique to carbonate aquifers; however, these vulnerable settings present a good test case for studies evaluating changes in pesticide concentrations over time. In these settings, detected compounds are numerous, typically have long half-lives, and are leachable; short half-life or nonleachable compounds generally are not observed (Loper et al., 2009; Lindsey et al., 2009; Lindsey et al., 2010).

Recovery adjusting pesticide concentrations makes it possible to compare the results from different sampling events if data on recoveries are available. Analyte recovery data are measured by analyzing water in quality-control samples that have been “spiked” with known amounts of target analytes. Bexfield (2008) noted that temporal changes in analytical recovery can adversely affect time–trend analysis of pesticide concentrations by introducing or masking trends in environmental concentrations that are caused by trends in performance of the analytical method rather than by trends in pesticide use or other environmental conditions. Models of pesticide recoveries of reagent-water spikes from long-term analytical performance datasets developed by the USGS National Water-Quality Assessment (NAWQA) program and the USGS National Water Quality Laboratory (NWQL) were used in the Bexfield (2008) study. Using reagent-water spike data, Bexfield (2008) adjusted pesticide concentrations (referred to hereafter as “recovery-adjusted concentrations”) reported by the NWQL to account for temporal changes in analytical method performance. The use of spikes in splits of well-water samples, or field matrix spiking, offers benefits over reagent-water spiking because analytical bias is measured in the sample and is more likely to account for matrix interferences (Martin et al., 2009). Previous investigations in Pennsylvania by the USGS, including those by Lietman (1997), Lindsey et al. (1998), Hainly et al. (2001), Fischer et al. (2004), and Loper et al. (2009), used unadjusted concentrations and did not use reagent water or matrix water recovery-adjusted pesticide concentrations when presenting or analyzing pesticide concentration data. Previous investigations outside of Pennsylvania have used matrix-based, recovery-adjusted concentrations to assess for changes in concentrations (H.M. Johnson et al., 2011; Saleh et al., 2011; Stackelberg et al., 2012; Sullivan et al., 2009). Recoveries differ among pesticides and may be less or greater than 100%, such that unadjusted concentrations may be less or greater than adjusted concentrations. Adjusted concentrations are deemed to be more representative of “actual” or true concentrations.

Groupwise temporal changes in water quality for two sampling times are commonly evaluated by using paired data statistical methods. Paired data are common in studies of pesticide occurrence, whether they be for two growing seasons from a research plot (Johnson et al., 2001) or for decadal samples from a regional aquifer (Mills et al., 2005; Bexfield, 2008; Debrewer et al., 2008). The traditional approaches to matched pair tests are the sign test and the Wilcoxon signed-rank test (Helsel and Hirsch, 2002) and have been used in previous studies of pesticides in groundwater. For example, the sign test was used in Bexfield (2008), and the Wilcoxon signed-rank test was used in Mills et al. (2005) and Debrewer et al. (2008). Both tests are available in numerous statistical packages. The Wilcoxon signed-rank test is commonly known to be more sensitive to changes because it includes an analysis of the magnitude of difference between paired observations, whereas the sign test only compares the number of increasing pairs with the number of decreasing pairs (Helsel and Hirsch, 2002). Neither of these tests includes zero-difference ties in the calculation of the test statistic. Omitting zero-difference ties in a data set can lead to an erroneously low probability value when assessing the difference between sample pairs (false positive) because the zero-difference tie cases are evidence of no change. Censored matched pairs (nondetections of a compound in one or both samples of a pair) are common in studies investigating pesticides in groundwater, and, in the case where both samples of a pair are nondetections, a zero-difference tie occurs (Table 1). Therefore, it is prudent that a paired data test is chosen that includes zero-difference ties in the computation of the test statistic.

The purpose of this study was to extend the methodology used in previous studies (Bexfield, 2008; Debrewer et al.,!2008) by using models of recoveries of matrix-water pesticide spikes from a long-term analytical performance dataset developed by USGS NWQL in recovery-adjusting concentrations and by applying statistical tests that account for zero-difference ties. A method to evaluate changes in matrix-based, recovery-adjusted pesticide concentrations is discussed by examining paired Prentice-Wilcoxon (PPW) (O’Brien and Fleming, 1987) results on data from two carbonate-rock aquifers in a predominantly agricultural area north of the Debrewer et al. (2008) study area. All chemical names of the pesticide compounds mentioned in this paper are listed in Supplemental Table S1.


<table>
<thead>
<tr>
<th>Pesticide compound</th>
<th>Number of sample pairs</th>
<th>Number of sample pairs</th>
<th>Number of sample pairs</th>
<th>Number of sample pairs</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Total</td>
<td>Detections in both samples of a sample pair (noncensored data only)</td>
<td>Detection in one sample and nondetection in one sample of a sample pair (noncensored and censored data)</td>
<td>Nondetections in both samples of a sample pair (censored data only, otherwise known as zero-difference ties)</td>
</tr>
<tr>
<td>Deethylatrazine</td>
<td>30</td>
<td>30</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Atrazine</td>
<td>30</td>
<td>30</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Metolachlor</td>
<td>30</td>
<td>6</td>
<td>6</td>
<td>18</td>
</tr>
<tr>
<td>Prometon</td>
<td>23†</td>
<td>13</td>
<td>4</td>
<td>6</td>
</tr>
<tr>
<td>Simazine</td>
<td>30</td>
<td>23</td>
<td>4</td>
<td>3</td>
</tr>
</tbody>
</table>

† The presence of prometon was verified but not quantified in one sample analyzed in 1993–1995 and six samples analyzed in 2008–2009; hence, there are 23 sample pairs to assess. All other frequently detected pesticide compounds had 30 sample pairs.
Materials and Methods

Study Area

The 30 wells sampled were completed in two carbonate-rock aquifers (Fig. 1), the Piedmont and Blue Ridge carbonate-rock aquifer in the Piedmont Physiographic Province (hereafter termed “Piedmont carbonate”) and the Valley and Ridge carbonate-rock aquifers in the Valley and Ridge Physiographic Province (including the Great Valley and Appalachian Mountain sections; hereafter termed “Valley and Ridge carbonate”). These settings in south-central and southeastern Pennsylvania were among the most vulnerable in the nation in a report by Gilliom et al. (2006) and were ranked the most vulnerable in an assessment of pesticide occurrence in Pennsylvania groundwater by Lindsey and Bickford (1999). Carbonate-rock aquifers, in general, are the most vulnerable aquifer type for pesticide and other anthropogenic sources of contamination of groundwater because of the potential for rapid movement of water from the surface to the groundwater system due to fractures in the bedrock caused by weathering and karst features such as sinkholes and caverns. Sinkhole density, in particular, relates to the occurrence of pesticides in groundwater from carbonate aquifers (Lindsey et al., 2010). The Valley and Ridge carbonate represents a high sinkhole-density system, and the Piedmont carbonate represents a medium sinkhole-density system with median densities of 28 and 4 sinkholes per 100 km², respectively (Lindsey et al., 2010). Flat topography and well drained soils are characteristic of the land overlying the carbonate-rock aquifers. Land use in the study area is predominantly agricultural row crop. Groundwater ages in the study area are likely a mixture of young and old water. G.C. Johnson et al. (2011) defines young water as having groundwater residence times less than 20 yr and old water as having groundwater residence times of 20 yr or more. Groundwater age data from 50 samples compiled by G.C. Johnson et al. (2011) in the Valley and Ridge carbonate aquifer indicated that approximately 20% of samples had groundwater residence times equal to 10 yr or less and approximately 80% of the samples had groundwater residence times equal to 20 yr or less.

Sample Collection and Analysis

One sample per site was collected in 1993–1995 as part of the USGS NAWQA program, and another sample from each site was collected in 2008–2009 as part of a cooperative effort between the USGS and the Pennsylvania Department of Agriculture. Existing domestic wells were selected that were completed in the unconfined aquifer and were generally less than 60 m (200 ft) deep and less than 20 yr old when sampled.
during the first sampling interval. Water level and physical properties of the water, such as specific conductance, pH, dissolved oxygen, and water temperature, were measured at the time of sample collection. For the 1993–1995 samples, major ions were measured to define the chemistry of the groundwater matrix. To allow for comparisons among the sampling sites and between sampling periods where applicable, all techniques and equipment used for sampling and analysis were in compliance with USGS protocols (USGS, 2006).

To provide comparable data for evaluating changes in pesticide concentrations, all 60 samples (30 samples collected in 1993–1995 and 30 samples collected in 2008–2009) were collected at the same time of year to approximate constant hydrologic conditions. Samples were collected during the summer when net recharge declines and typically after any application of pre-emergent pesticides. Both sets of samples were analyzed by the USGS NWQL for 47 pesticides or pesticide degradates using gas chromatography/mass spectrometry with selected ion monitoring (Zaugg et al., 1995). As part of routine quality-assurance measures, study-specific samples of pesticide-spiked matrix waters were submitted to the NWQL, contributing to a long-term analytical performance dataset, at a rate of about 5% of the total number of samples collected (3 of 30 wells sampled 1993–1995 and 1 of 30 wells sampled 2008–2009). The 1993–1995 matrix spike data from this study were included in the USGS NWQL analytical performance datasets from multiple studies for national assessment and models for 999 NAWQA groundwater samples in Martin and Eberle (2011).

Recovery Adjustment

Martin et al. (2009) found that temporal changes in recovery were similar for laboratory reagent-water spikes and matrix spikes for most pesticides but that models of recoveries based on matrix spikes were expected to more closely match the matrix of environmental water samples. The work of Martin et al. (2009) on samples analyzed by the NWQL from 1992 to 2006 was recently updated for samples collected through 2010 and models for recovery of 44 pesticides and eight pesticide degradates, based on robust, locally weighted scatterplot smooths (LOWESS smooths) of matrix spikes, are now available for groundwater (Martin and Eberle, 2011).

For pesticides, recovery is the measured amount of pesticide in the spiked QC sample expressed as a percentage of the amount spiked, ideally 100%. The models of recovery can be used to adjust concentrations of pesticides measured in groundwater to 100% recovery to compensate for temporal changes in the performance (bias) of the analytical method.

For this study, pesticide concentrations of the five frequently detected pesticide compounds (atrazine, metolachlor, simazine, prometon, and DEA) from the 1993–1995 and 2008–2009 samples were adjusted to 100% recovery using modeled values for recovery in groundwater based on 1992–2010 matrix spike data for the day of sample collection from Appendix 4 in Martin and Eberle (2011). This approach assumes that the modeled recovery is the most accurate correction factor for the individual sample. The assumption allows a cost-effective approach in situations where it would be prohibitively expensive to spike each sample and attempt to document recovery and possible differences in recovery related to the specific matrix of each sample. For each environmental sample having a quantified detection of an individual pesticide compound, the unadjusted concentration was multiplied by 100 and divided by the modeled recovery to adjust the concentration to 100% recovery. Non-detections were reported as censored concentrations at the maximum long-term method detection level (LT-MDL) for that individual pesticide compound as reported in Martin and Eberle (2011).

Statistical Analysis

Some analytes in the dataset were considered highly censored from a statistical standpoint because of a high percentage of concentrations that were below reporting levels (commonly referred to as “nondetects”). Multiple reporting levels for individual pesticide compounds were common over time, and reporting levels varied among the five frequently detected compounds analyzed by the NWQL between the first and second sampling intervals (1993–1995 and 2008–2009). The dataset also included estimated concentrations for pesticides that were reliably detected but had a degree of uncertainty associated with the quantitative measurement. For this study, the hierarchy of the dataset was preserved by considering all censored pesticide values as equivalent to the maximum LT-MDL reported by Martin and Eberle (2011) for the individual pesticides and all estimated values as quantified detections higher than the censored pesticide values, unless otherwise noted.

The nonparametric PPW test was used to test for temporal changes in pesticide concentrations (O’Brien and Fleming, 1987; Helsel, 2005) in the paired observation data from the 30 wells sampled twice. The PPW test is a score test that can handle censored matched pairs (nondetects of a compound in one or both sampling intervals). Scores were computed for each observation as a measure of the position of the observation in the dataset. Differences between scores for 1993–1995 pesticide concentrations compared with scores for 2008–2009 pesticide concentrations at each site were computed to determine whether the sum of the differences for the entire dataset was significantly different from zero at the 95% confidence level (Helsel, 2005). Pesticide concentrations were censored to a common assessment level (the maximum LT-MDL for individual pesticides) before using the PPW test as recommended by O’Brien and Fleming (1987). Not censoring to a common assessment level in cases where one event is consistently censored at a different value than the other results in a probability value that is artificially low based solely on the differences in censoring levels. For each pesticide, any concentration (reported or estimated) falling below the maximum LT-MDL was classified as a nondetect. The PPW test was only performed on the recovery-adjusted concentrations for the five frequently detected pesticide compounds (DEA, atrazine, simazine, metolachlor, and prometon) using Minitab software (version 15) and a macro by Helsel (2005). As part of the PPW macro, an estimate of the median difference between groups for the frequently detected pesticides was determined using the Turnbull method for censored observations (Helsel, 2005).

Results and Discussion

Detection Frequency

Of the 47 pesticide compounds analyzed by gas chromatography/mass spectrometry during both sampling intervals (1993–1995 and 2008–2009), the pesticides most
frequently detected in samples of groundwater from the Valley and Ridge and the Piedmont carbonate aquifers were the triazine herbicides atrazine, simazine, and prometon; the acetanilide herbicide metolachlor; and an atrazine degradate (DEA). Atrazine and DEA had the highest detection frequencies, with concentrations above reporting levels in all 60 samples analyzed (Supplemental Table S2). Detections of the five frequently detected compounds were widespread in the Piedmont and Valley and Ridge carbonate aquifers, but all pesticide concentrations (unadjusted) were less than USEPA maximum contaminant levels (USEPA, 2011; Supplemental Table S2). Because pesticide concentrations reported by most laboratories are unadjusted concentrations, pesticides typically are regulated based on those concentrations.

Other studies have indicated that the Piedmont and Valley and Ridge carbonate aquifers are particularly vulnerable to pesticide contamination. For example, a study by Lindsey et al. (2009) in selected carbonate aquifers in the United States reported that the Piedmont and the Valley and Ridge carbonate aquifers had the most frequent detections of pesticides, along with the Biscayne aquifer in Florida, out of 12 aquifers assessed. The Piedmont and the Valley and Ridge carbonate aquifers also had the highest percentages of agricultural land use and the highest reported use of atrazine and metolachlor, along with the Silurian Devonian/Upper carbonate aquifer in eastern Iowa and southern Minnesota (Lindsey et al., 2009).

Detection frequencies for common pesticides in water samples from sites in the Piedmont and the Valley and Ridge carbonate aquifers (atrazine, deethylatrazine, metolachlor, prometon, and simazine) generally are consistent with the hypothesis that high-use pesticides with high leaching potential will be detected more frequently than low-use pesticides with low leaching potential. For example, in a study by G.C. Johnson et al. (2011), chlorpyrifos was among the 15 high-use pesticides within the Valley and Ridge carbonate, but it has a low leaching potential (Lindsey and Bickford, 1999; G.C. Johnson et al., 2011) and is among the most frequently detected pesticides in this study (Supplemental Table S1).

On the other hand, atrazine, simazine, and metolachlor, which also are widely used in the Valley and Ridge (G.C. Johnson et al., 2011), have high leaching potential (Lindsey and Bickford, 1999; G.C. Johnson et al., 2011) and are among the most frequently detected pesticides in this study (Supplemental Table S1).

Every sample in the study area had detections of multiple pesticide compounds. In the Valley and Ridge carbonate aquifer, 12 of 20 samples (1993–1995) and 15 of 20 samples (2008–2009) had five or more pesticides detected, and 5 of 10 samples (1993–1995) and 3 of 10 samples (2008–2009) had five or more pesticides detected in the Piedmont carbonate aquifer (Fig. 2). Finding multiple pesticides in samples from unconfined carbonate aquifers overlain by agricultural land use in this study is consistent with findings from other studies (Lindsey et al., 2009; G.C. Johnson et al., 2011). Dissolved oxygen measurements indicate that both aquifers in the current study are characterized by oxygenated groundwater (Table 2), a finding that has been associated with the occurrence of multiple pesticides in this

Fig. 2. Groundwater samples with one or two, three or four, and five or more pesticide compounds detected in the Piedmont and Valley and Ridge carbonate-rock aquifers in 1993–1995 and 2008–2009. Results are based on analysis of 47 pesticide compounds.

Table 2. Water-quality characteristics and major ion concentrations in groundwater from wells sampled in two carbonate-rock aquifers, Pennsylvania.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Characteristics</td>
<td>Range of measured values†</td>
<td>Range of measured values†</td>
</tr>
<tr>
<td>Water temperature, °C</td>
<td>11.5–14.0</td>
<td>10.5–15.0</td>
</tr>
<tr>
<td>Specific conductance, μS cm⁻¹</td>
<td>340–960</td>
<td>438–970</td>
</tr>
<tr>
<td>Dissolved oxygen, mg L⁻¹</td>
<td>0.9–10.6</td>
<td>6.1–10.0</td>
</tr>
<tr>
<td>pH</td>
<td>7.0–7.5</td>
<td>7.1–7.7</td>
</tr>
<tr>
<td>Major ions‡</td>
<td>Concentration range, mg L⁻¹</td>
<td>Concentration range, mg L⁻¹</td>
</tr>
<tr>
<td>Calcium</td>
<td>39.0–140</td>
<td>48–110</td>
</tr>
<tr>
<td>Magnesium</td>
<td>5.4–43</td>
<td>7.2–48</td>
</tr>
<tr>
<td>Sodium</td>
<td>3.4–15</td>
<td>1.6–54</td>
</tr>
<tr>
<td>Potassium</td>
<td>0.6–5.1</td>
<td>0.8–5.1</td>
</tr>
<tr>
<td>Bicarbonate</td>
<td>112–316</td>
<td>212–400</td>
</tr>
<tr>
<td>Chloride</td>
<td>10.0–72</td>
<td>6.5–64</td>
</tr>
<tr>
<td>Sulfate</td>
<td>10.0–130</td>
<td>8.2–33</td>
</tr>
<tr>
<td>Silica (as SiO₂)</td>
<td>7.0–13</td>
<td>5.1–13</td>
</tr>
<tr>
<td>Nitrate plus nitrite (as N)</td>
<td>2.0–25</td>
<td>5.5–18</td>
</tr>
</tbody>
</table>

† Range observed during dates sampled.
‡ Major constituents and nutrients analyzed by U.S. Geological Survey National Water Quality Laboratory in samples filtered through a membrane with 0.45-µm pore size.
setting (G.C. Johnson et al., 2011). Near neutral pH and calcium-bicarbonate water types predominate (Table 2). In addition to the presence of pesticides, most groundwater samples from the two aquifers have elevated nitrate concentrations (>10 mg L⁻¹ as N) characteristic of these agricultural settings.

**Changes in Concentrations**

For the frequently detected pesticides, differences in paired observation data (concentrations from the 30 sites sampled one time in 1993–1995 and concentrations from each of the same wells sampled one time in 2008–2009) were examined. The PPW test identified significant (95% confidence level) decreases in the recovery-adjusted concentrations of atrazine, DEA, simazine, and prometon in 1993–1995 samples as a whole compared with 2008–2009 samples in the Piedmont and Valley and Ridge carbonate-rock aquifers (Fig. 3). Metolachlor concentrations were not significantly different between 1993 and 2009 (z-score = 0.109; P = 0.913). Specific conductance was not significantly different in 1993–1995 samples compared with 2008–2009 samples (z-score = 1.54; P = 0.124), a gross indication that chemically similar groundwaters were sampled during the two periods. Water level (depth to water) in wells was not significantly different in 1993–1995 compared with 2008–2009 (z-score = 1.64; P = 0.101), a gross indication that hydrologic conditions were similar during the two periods.

Results for recovery-adjusted pesticide concentrations for the five frequently detected pesticides were compared with results that would have been obtained using unadjusted concentrations (Fig. 4). Results of the PPW test using unadjusted pesticide concentrations indicated significant decreases in concentration for atrazine and simazine, significant increases in concentration for DEA, and no significant difference for metolachlor or prometon concentrations from 1993 to 2009 at the 95% confidence level. Conclusions using unadjusted concentrations (i.e., no change in concentration for prometon and the direction of concentration change for DEA) differed from conclusions using recovery-adjusted concentrations. Large differences in recoveries between periods, such as those for DEA in particular, that were relatively low in the 1993–1995 samples compared with 2008–2009 (Supplemental Table S3) might be expected to affect results of the PPW tests for change. Unadjusted and recovery-adjusted concentrations for all 60 observations for this study are shown in Supplemental Table S3, and the study-specific recoveries (wells MF 401, CU 919, and LN 2027) are shown in Martin and Eberle (2011). In general, the study-specific recoveries were within 20% or less of the modeled recoveries for the five pesticides and were commonly higher than the modeled recoveries, perhaps as a result of matrix effects such as occurrence of multiple pesticides. A bias, if uniform, in modeled recoveries relative to the study specific recoveries would not adversely affect the time–trend analysis (Martin et al., 2009).

Similar to what Bexfield (2008) reported, differences in results of statistical analysis for changes in recovery-adjusted concentrations compared with unadjusted concentrations showed that changes not accounted for in analytical recovery can mask a true change, misidentify a change when no true change exists, or assign a direction opposite of the true change in concentration that was primarily the result of the performance of the analytical method. Similarity of statistical test results for changes in atrazine, metolachlor, and simazine concentrations, regardless of whether recovery-adjusted or unadjusted data were used, indicate that more confidence can be placed in concentration changes observed for those pesticides than for DEA or prometon for this study.

The pattern of decline in concentrations observed for atrazine, DEA, simazine, and prometon concentrations for this study was consistent with results of other groundwater studies in similar settings. Kolpin et al. (1997) reported decreasing unadjusted concentrations of atrazine in groundwater samples collected from selected municipal wells in agricultural areas of Iowa over time (1982–1995). Conversely, Kolpin et al. (1997) reported increasing unadjusted concentrations of metolachlor in the Iowa wells for the period 1982–1995. A study of groundwater quality in agricultural areas of the mid-Atlantic region, including seven sites in carbonate-rock aquifers of the Great Valley Potomac River basin in Pennsylvania (south of the area on which this article is based), reported significantly decreasing recovery-adjusted concentrations of atrazine, DEA, and prometon in 2002 compared with 1993 (Debrewer et al., 2008). A study of nine pairs of pesticide results for groundwater in agricultural areas in the Lower Susquehanna River basin in Pennsylvania reported increasing unadjusted concentrations of atrazine and decreasing unadjusted concentrations of simazine in groundwater samples collected from selected municipal wells in agricultural areas of Iowa over time (1982–1995). Conversely, similar to what Bexfield (2008) reported, differences in results of statistical analysis for changes in recovery-adjusted concentrations compared with unadjusted concentrations showed that changes not accounted for in analytical recovery can mask a true change, misidentify a change when no true change exists, or assign a direction opposite of the true change in concentration that was primarily the result of the performance of the analytical method. Similarity of statistical test results for changes in atrazine, metolachlor, and simazine concentrations, regardless of whether recovery-adjusted or unadjusted data were used, indicate that more confidence can be placed in concentration changes observed for those pesticides than for DEA or prometon for this study.

The pattern of decline in concentrations observed for atrazine, DEA, simazine, and prometon concentrations for this study was consistent with results of other groundwater studies in similar settings. Kolpin et al. (1997) reported decreasing unadjusted concentrations of atrazine in groundwater samples collected from selected municipal wells in agricultural areas of Iowa over time (1982–1995). Conversely, Kolpin et al. (1997) reported increasing unadjusted concentrations of metolachlor in the Iowa wells for the period 1982–1995. A study of groundwater quality in agricultural areas of the mid-Atlantic region, including seven sites in carbonate-rock aquifers of the Great Valley Potomac River basin in Pennsylvania (south of the area on which this article is based), reported significantly decreasing recovery-adjusted concentrations of atrazine, DEA, and prometon in 2002 compared with 1993 (Debrewer et al., 2008). A study of nine pairs of pesticide results for groundwater in agricultural areas in the Lower Susquehanna River basin in Pennsylvania reported increasing unadjusted concentrations of atrazine and decreasing unadjusted concentrations of simazine in groundwater samples collected from selected municipal wells in agricultural areas of Iowa over time (1982–1995). Conversely, similar to what Bexfield (2008) reported, differences in results of statistical analysis for changes in recovery-adjusted concentrations compared with unadjusted concentrations showed that changes not accounted for in analytical recovery can mask a true change, misidentify a change when no true change exists, or assign a direction opposite of the true change in concentration that was primarily the result of the performance of the analytical method. Similarity of statistical test results for changes in atrazine, metolachlor, and simazine concentrations, regardless of whether recovery-adjusted or unadjusted data were used, indicate that more confidence can be placed in concentration changes observed for those pesticides than for DEA or prometon for this study.

![Fig. 3. Paired differences in concentrations 2008–2009 minus 1993–1995 for 30 wells for deethylatrazine, atrazine, metolachlor, prometon, and simazine in the Piedmont and Valley and Ridge carbonate-rock aquifers in Pennsylvania. z-score, median difference (Turnbull method), and P value indicate results of the paired Prentice-Wilcoxon test (Helsel, 2005).](image-url)
Fig. 4. (A) Unadjusted and (B) recovery-adjusted concentrations for selected pesticide compounds in 30 wells sampled one time in 1993–1995 and one time again in 2008–2009 in predominantly agricultural areas in the Piedmont and Valley and Ridge carbonate-rock aquifers in Pennsylvania.

the Valley and Ridge and Piedmont carbonate-rock aquifers in Pennsylvania reported significantly decreasing unadjusted atrazine concentrations in 2003 compared with 1993–1995 (Loper et al., 2009). Five of the nine wells analyzed by Loper et al. (2009) were used for the analysis herein.

**Conclusions**

Studies evaluating temporal changes in paired observations of pesticide concentration in groundwater should incorporate (i) a study design where hydrologic conditions are similar during sampling intervals, (ii) consistent methods to extract and analyze the samples using the same laboratory throughout the study, (iii) systematic submission of pesticide spikes in matrix water to allow accounting for differences in analytical recoveries of pesticides and concentration adjustment to remove variability solely from changes in performance of the analyzing laboratory, (iv) a knowledge of geochemical characteristics of the samples in the matrix spike recovery database so that they are as similar as possible to the matrix of groundwater samples subject to adjustment of pesticide concentration with modeled recoveries, and (v) a statistical test appropriate for the dataset, such as the PPW test, that is most useful for studies analyzing the magnitude of change on datasets that include zero-difference ties (nondetections of a compound in both samples of a pair) and multiple censored data.

Differentiating sample matrix effects from laboratory performance variations needs more attention. More studies are needed to determine if recovery models derived from a laboratory performance dataset from multiple studies for national assessment are appropriate for making time- and study-specific adjustments for analyte recovery.

**Acknowledgments**

This study was conducted by the U.S. Geological Survey in cooperation with the Pennsylvania Department of Agriculture. The authors thank Donald Gilbert and David Scott of the Pennsylvania Department of Agriculture for assistance and support throughout the study and the following colleagues from the U.S. Geological Survey who provided technical support or field assistance during this study: Jack Barbash, Jeff Chaplin, Justin Kudon, Bruce Lindsey, Connie Loper, Jeff Martin, Gail Thelin, Naomi Nakagaki, and Linda Zarr. Bruce Lindsey and three anonymous reviewers provided comments and suggestions that improved the manuscript.

**References**


