AN EVALUATION
of the
STATISTICAL POWER OF VOLUNTEER-
GENERATED DATA COMPILED
by the
UPPER MERRIMACK MONITORING PROGRAM

Stephen C. Landry
Michele L. Tremblay
principal investigators

Upper Merrimack Monitoring Program
a citizen watershed monitoring program
An Evaluation of the Statistical Power of Volunteer-Generated Data Compiled by the Upper Merrimack Monitoring Program
February 2002

Principal Investigators
Stephen C. Landry, Sampling Supervisor, Quality Control/Quality Assurance Officer
Michele L. Tremblay, Program Director

Abstract
There are increasing numbers of groups in New Hampshire who work with volunteers to monitor water quality. Some, such as agencies and scientists, question the validity of the data that volunteer groups produce. Unfortunately, citizen volunteers often lack specific, scientific knowledge and training that would make their efforts more relevant and effective.

The Upper Merrimack Monitoring Program (UMMP) began a volunteer river assessment effort in 1995 that trained volunteers to collect physical, chemical and biological data. The UMMP field and laboratory procedures comply with a Quality Assurance/Quality Control Plan approved by the US Environmental Protection Agency (EPA) with methodologies that parallel those employed by state biologists. In an effort to test the statistical power of volunteer data generated by the UMMP, replicate samples were analyzed by volunteers and professional taxonomists from an independent contract laboratory (EcoAnalysts, Inc.) respectively. The Family level taxonomy returned from the contract lab and the identifications generated by UMMP volunteers were used independently to calculate three biological metrics for each site. A total of eleven sites were audited in this manner with a total of twenty-two replicates analyzed. The Wilcoxon Signed-Rank Test for Paired Data was applied to each set of replicate biometric scores generated from UMMP and EcoAnalysts taxonomy respectively. A null hypothesis was formed predicting that biometric scores generated with UMMP taxonomy data would not be statistically different from biometric scores generated with EcoAnalysts taxonomy data. Results from the Wilcoxon Signed-Rank Test for Paired Data indicated that there was not enough evidence (the test statistic value was not small enough to disprove the null hypothesis) to show that the two groups (UMMP and EcoAnalysts taxonomy) are different for both EPT and Family Biotic Index Scores.

Results from the Wilcoxon Signed-Rank Test for Family Richness scores produced evidence sufficient enough to reject the null hypothesis concluding that scores generated from EcoAnalysts data are statistically different (larger) from those scores generated from UMMP data. Although this result appears to weaken the position of the volunteer generated data, upon closer inspection, it becomes evident that the Family Richness biometric scores are influenced not only by family level taxonomy expertise but also by variables during collection and laboratory sorting. In fact, the rejection of the null hypothesis for the Family Richness scores strengthens the validity of UMMP taxonomy in the lab since the EPT and Family Biotic Index scores sometimes rely upon the correct identification of a single organism in order to record a score for particular biometrics.

The fact that the UMMP generated biometric scores for EPT and the Family Biotic Index were not statistically different from those calculated with EcoAnalysts data demonstrates the validity of UMMP generated taxonomy in the laboratory. Therefore, the volunteer data was indistinguishable from professional level data during the three-year reporting period from 1995 to 1997 as presented in the program’s State of the Upper Merrimack 1995-1997, a river quality report.
Background
The Upper Merrimack Monitoring Program diagnoses river quality from the Upper Merrimack (Franklin to Bow), Pemigewasset, Winnipesaukee, and Contoocook Rivers using E. coli bacteria, field chemistry, habitat assessments, and benthic macroinvertebrates. The volunteer-staffed program began in 1995, with a cooperative agreement between the Upper Merrimack River Local Advisory Committee, Merrimack River Watershed Council, and the NH Department of Environmental Services (DES). During that time, the UMMP sampled seven sites. In 1996, the program was expanded to include a total of eleven sites. In 1996, the UMMP received the first of two grants from the New England Interstate Water Pollution Control Commission. These grants allowed the UMMP sustainability and independence by funding the purchase of equipment and support materials. The second grant provided support to expand the UMMP's strong educational component.

Hundreds of volunteers from all over the region have spent countless hours participating in the UMMP. River conservationists, municipal officials, anglers, teachers, students, and many others have assisted the UMMP by collecting water samples and benthic macroinvertebrate specimens, analyzing macroinvertebrates, and performing other field tasks. Their willingness to learn and spend their free time is at the core of the UMMP's success—making it one of the premier and most ambitious volunteer water quality monitoring programs in the country.

The UMMP collects and generates data by adhering to EPA-approved Quality Assurance/Quality Control Plan. Adherence to the Plan ensures consistency in the collection and analysis of UMMP data but also validates the information generated for incorporation by State and Federal agencies.

Sample Parameters and Methodologies
The following sample parameters have been employed by the Upper Merrimack Monitoring Program:

- E. coli bacteria
- Field chemistry
- Habitat Assessment
- Biomonitoring with benthic macroinvertebrates

E. coli samples are collected using WHIRL-PAKs™ and are processed by the Franklin Waste Water Treatment Facility. The NH Department of Environmental Services Laboratory Services Unit processes some samples as a random quality assurance/quality control measure. Field chemistry samples were collected and analyzed using a Hydrolab™ borrowed from the DES. Habitat assessments were completed by UMMP volunteers at each site by following Rapid Bioassessment Protocol (RBP) guidelines identical to those utilized by EPA as well as DES biologists.

The Upper Merrimack Monitoring Program focuses much of its effort on biomonitoring with benthic macroinvertebrates. UMMP volunteers deploy an array of three rock baskets at each monitoring site in order to analyze benthic macroinvertebrate community composition. These rock baskets or “artificial substrates” are deployed for a seven-week colonization period which conforms to EPA-RBP protocols. At the conclusion of the colonization period, volunteers are trained in proper retrieval methods to ensure consistent sample handling at all UMMP sites. Benthic macroinvertebrates are removed from the substrates and preserved for future sorting and identification by UMMP volunteers.
Quality Assurance/Quality Control
The Upper Merrimack Monitoring Program follows benthic macroinvertebrate field and laboratory methodologies outlined by the River Watch Network Benthic Macroinvertebrate Monitoring Manual which are referenced in the UMMP Quality Assurance/Quality Control Plan. The collection of E. coli samples by volunteers follow the sample collection guidelines outlined by Standard Methods. Deployment of a Hydrolab® multi-parameter water quality monitor and the compilation of habitat assessment forms comply with DES Biomonitoring Program protocols which are consistent with methodologies followed by EPA personnel.

Once invertebrate samples have been collected and preserved in the field, UMMP volunteers conduct sub-sampling and taxonomic identification to the family level. An annual series of macroinvertebrate processing and identification workshops are scheduled that are open to all UMMP volunteers and members of the public. The workshops, popularly known as “Bug Nights” provide volunteers with the training necessary to perform sub-sampling and family level identifications upon macroinvertebrates collected by the UMMP.

The UMMP provides reference materials necessary for volunteers to accurately identify macroinvertebrates to the family level. Widely accepted reference texts, such as McCafferty, Wiggins, Merritt, and Cummins, are utilized in this process. In an effort to streamline the identification process, and to make the standard taxonomic keys less daunting for volunteers unaccustomed to using these texts, the UMMP produced Macroinvertebrates of the Upper Merrimack, an annotated picture key. This key contains only the macroinvertebrate families identified within the upper Merrimack watershed, which allows volunteers to make positive identifications for the most commonly encountered organisms without having to negotiate through lengthy dichotomous keys. The picture key is updated annually as new families are identified.

A series of inter-laboratory audits are performed on a regular basis during the identification workshops to ensure that high quality data is being produced. Another UMMP leader, volunteer, or professional biologist performs an initial audit once a replicate is “sub-sampled.” The audit consists of a separate individual examining the contents of the sorting tray to make sure no organisms have been overlooked, a minimum of three out of twelve grids have been “picked” and that the sub-sample vial contains the minimum 100 organisms. The secondary audit typically involves the UMMP Director, Sampling Supervisor, or a professional biologist confirming difficult identifications and verifying completed vials that volunteers have processed.

The UMMP is proud that so many volunteers return each year to participate in “Bug Nights.” The majority of volunteers demonstrate meticulous sample handling, attention to detail, and collaboration with workshop leaders and attendees to ensure accurate family level identifications. Although the intra-program quality assurance/quality control measures are rigorously maintained by the UMMP, it was decided that an outside, professional-level audit was necessary to further validate the volunteer generated data.

The publication by UMMP of the State of the Upper Merrimack 1995-1997 documents three years of physical, chemical, and biological river quality within the upper Merrimack watershed. This report also demonstrated the ability of a volunteer organization to generate data that was acceptable for state and federal agencies to consider in their decision-making processes. Although, the incorporation of the UMMP data into the 305(b) Report to Congress is considered a huge success, there remain some that are unwilling to accept data generated by volunteers.

One of the core elements of the UMMP is to validate data generated by volunteer monitoring organizations. During the 1996 and 1997 monitoring seasons, several replicate
samples were sent to an independent, professional laboratory for family level identifications of invertebrates collected by UMMP volunteers. Replicates from a single site were separated in order to have volunteer identification conducted on “Replicate A” while professional entomologists from EcoAnalysts analyzed “Replicate B”. Once data was returned to the UMMP from EcoAnalysts, the taxa list from the UMMP volunteers could be compared to the taxa list submitted by the lab.

It was immediately evident that conducting a simple comparison of family richness between the UMMP replicate and the replicate analyzed by EcoAnalysts did not provide the comparative analyses necessary for validating the volunteer efforts. Although three rock baskets are deployed on the river bottom in an array at a given site, the small differences in flow, sediment deposition, and position can yield a dramatically different population from basket to basket. These differences will affect organism densities, community structure, and function as the colonization potential of each basket develops independently over the seven-week colonization period.

The family identifications returned to UMMP from EcoAnalysts did provide a degree of verification in that there were no significant “surprises” when volunteer lists were compared to the professional laboratory lists. In many instances, EcoAnalysts identified more families at a given site when compared to the replicates processed by UMMP volunteers. A closer examination of the family richness lists indicates that the replicate processed by EcoAnalysts had a much higher organism density when compared to the replicate processed by UMMP. This condition will usually lead to additional taxa being present within a given sample or replicate.

In order to test the statistical validity and power of the UMMP generated data; it is necessary to compare biometric scores calculated with UMMP and EcoAnalysts data to determine if they are statistically different. This process will not only validate the UMMP generated data but provide a statistical foundation for the data presented in the State of the Upper Merrimack 1995-1997.

**Statistical Power Analysis**

The UMMP utilizes family-level macroinvertebrate data to calculate a series of biological metrics or biometrics in order to assess river quality. Given the high degree of variability in family richness values among replicates at a single site, a statistical comparison of calculated biometric scores generated with UMMP data versus EcoAnalysts data provides a much more valid determination of similarity.

Although replicate samples are retrieved from the same site, independently, they are considered random samples since the potential for colonization varies so dramatically. In order to compare two independently analyzed replicates and the subsequent biometric scores to decide if they are significantly different, a Wilcoxon Signed-Rank Test for Paired Data can be performed. This test assumes that x and y are continuous, random variables and in this case x and y represent biometric scores generated with UMMP and EcoAnalysts data respectively.

The Wilcoxon Signed-Rank Test for Paired Data was applied to a set of twenty-two biometric values for family richness. The family richness metric was selected for comparative analysis because out of all the biometrics employed by the UMMP; family richness has the highest potential for variability from basket to basket. Eleven values (x) were calculated from UMMP-generated family level identifications with matching replicates and their family level identifications provided by EcoAnalysts (y). The Wilcoxon Signed-Rank Test was applied to the data set to determine if the family richness biometric scores were significantly different. The null
hypothesis for all three biometric analyses was that the UMMP biometric scores \((x)\) were the same as the biometric scores calculated with EcoAnalysts taxonomic information \((y)\).

Table 1 clearly illustrates the range of family richness values encountered between replicates from a single site. The difference of family richness values ranged from zero to seven when comparing replicates from the same site. This is purely a function of organism density and variation among individual baskets at a given site. There is also a strong degree of randomness introduced into this biometric since sample vials are selected by volunteers during “Bug Nights” where the family level taxonomy is performed each winter. The remaining vials that are uncompleted get sent to EcoAnalysts for professional level analyses and comparison.

<table>
<thead>
<tr>
<th>UMMP</th>
<th>EcoAnalysts</th>
<th>Diff</th>
<th>Diff</th>
<th>Rank</th>
<th>Signed Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
<td>12</td>
<td>-5</td>
<td>5</td>
<td>9.5</td>
<td>-9.5</td>
</tr>
<tr>
<td>10</td>
<td>12</td>
<td>-2</td>
<td>2</td>
<td>6.5</td>
<td>-6.5</td>
</tr>
<tr>
<td>9</td>
<td>10</td>
<td>-1</td>
<td>1</td>
<td>3.5</td>
<td>-3.5</td>
</tr>
<tr>
<td>11</td>
<td>16</td>
<td>-5</td>
<td>5</td>
<td>9.5</td>
<td>-9.5</td>
</tr>
<tr>
<td>16</td>
<td>18</td>
<td>-2</td>
<td>2</td>
<td>6.5</td>
<td>-6.5</td>
</tr>
<tr>
<td>9</td>
<td>10</td>
<td>-1</td>
<td>1</td>
<td>3.5</td>
<td>-3.5</td>
</tr>
<tr>
<td>16</td>
<td>19</td>
<td>-3</td>
<td>3</td>
<td>8</td>
<td>-8</td>
</tr>
<tr>
<td>9</td>
<td>16</td>
<td>-7</td>
<td>7</td>
<td>10</td>
<td>-10</td>
</tr>
<tr>
<td>15</td>
<td>15</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>7</td>
<td>8</td>
<td>-1</td>
<td>1</td>
<td>3.5</td>
<td>-3.5</td>
</tr>
<tr>
<td>18</td>
<td>17</td>
<td>1</td>
<td>1</td>
<td>3.5</td>
<td>3.5</td>
</tr>
</tbody>
</table>

The “P” value (from Table XI in Appendix B of Statistical Methods in the Biological and Health Sciences) of the observed value of the statistic is less than .01. Since this value is so small, the null hypothesis is rejected. It is concluded that the EcoAnalysts samples are different than UMMP samples—specifically because they’re larger. Although the Wilcoxon Signed-Rank test for paired data rejects the null hypothesis for the Family Richness Scores (the null hypothesis stating that UMMP generated biometric scores will not be statistically different from those scores generated with EcoAnalysts data), this cannot be viewed as a failure of volunteer generated data. In hindsight, the inclusion of the Family Richness Scores to test the statistical power of volunteer versus professional level taxonomy along with the biometric scores derived from that information may be suspect due to the extremely high degree of randomness associated with invertebrate sample selection and processing during the volunteer identification sessions.

During the three years of data collection, processing and analyses, UMMP volunteers subsampled and identified macroinvertebrates at “Bug-Nights” that were regularly scheduled from 6:00 PM to 9:00 PM. Volunteers are allowed to select samples at random for subsampling and family level identifications during these sessions. Given that there is a three-hour limit to the work that can be completed on a given “Bug Night,” volunteers may have selected samples that appeared to have a “manageable” amount of organisms and debris within a particular jar. This would explain why the organism densities and subsequent family richness values were consistently higher within the samples delivered to EcoAnalysts for processing. In future laboratory audits, UMMP volunteers will perform the Family level identifications on specific samples, which will then be shipped out to an independent laboratory for taxonomic
verification. The Family Richness biometric will then be calculated using the respective identification lists completed by UMMP volunteers as well as from the independent laboratory. This will provide a much more valid statistical analysis for this particular biometric. Figure 1 graphically presents the family richness values calculated with UMMP and EcoAnalysts data.

Figure 1: UMMP vs. EcoAnalysts Generated Family Richness Scores

The Wilcoxon Signed Rank Test for paired data was also applied to the biometric scores calculated for Percent EPT as derived from data generated by the UMMP and EcoAnalysts respectively. Percent EPT or the percentage of Ephemeroptera (Mayfly nymphs), Plecoptera (Stonefly nymphs), and Trichoptera (Caddisfly nymphs) taxa is a common biometric that quantifies the percentage of the invertebrate community comprised of the number of families found within these three orders. A healthy benthic community should have a relatively high percentage of EPT families present and therefore a high Percent EPT value.

The decision to include this particular biometric in the statistical power analysis was two-fold: first, the mis-identification or lack of recognition of one of these key families can alter the biometric scores significantly. Secondly, the family identifications of particular Mayfly, Stonefly, and even some Caddisflies can be extremely challenging. In addition to these two specific qualifies, the EPT biometric is extremely important for UMMP data reporting since Caddisfly families dominate the invertebrate community encountered at virtually all UMMP sites.

Table 2 presents the results of the Wilcoxon Signed Rank Test for Family EPT scores generated with UMMP and EcoAnalysts data respectively.
Table 2: Family EPT Scores

<table>
<thead>
<tr>
<th>UMMP</th>
<th>EcoAnalysts</th>
<th>Diff</th>
<th></th>
<th>Diff</th>
<th>Rank</th>
<th>Signed Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>.98</td>
<td>.83</td>
<td>.15</td>
<td>.15</td>
<td>10</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>.72</td>
<td>.86</td>
<td>-.14</td>
<td>.14</td>
<td>9</td>
<td>-9</td>
<td></td>
</tr>
<tr>
<td>.89</td>
<td>.85</td>
<td>.04</td>
<td>.04</td>
<td>3</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>.59</td>
<td>.58</td>
<td>.01</td>
<td>.01</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>.80</td>
<td>.89</td>
<td>-.09</td>
<td>.09</td>
<td>7</td>
<td>-7</td>
<td></td>
</tr>
<tr>
<td>.98</td>
<td>.91</td>
<td>.07</td>
<td>.07</td>
<td>6</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>.88</td>
<td>.92</td>
<td>-.04</td>
<td>.04</td>
<td>3</td>
<td>-3</td>
<td></td>
</tr>
<tr>
<td>.47</td>
<td>.79</td>
<td>.18</td>
<td>.18</td>
<td>11</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td>.93</td>
<td>.89</td>
<td>.04</td>
<td>.04</td>
<td>3</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>.17</td>
<td>.21</td>
<td>-.10</td>
<td>.10</td>
<td>8</td>
<td>-8</td>
<td></td>
</tr>
<tr>
<td>.42</td>
<td>.48</td>
<td>-.06</td>
<td>.06</td>
<td>5</td>
<td>-5</td>
<td></td>
</tr>
</tbody>
</table>

\[ W = |W_+| = \sum \text{all positive ranks } R_i = 10 + 3 + 1 + 6 + 11 + 3 = 34 \]

\[ W = |W_-| = \sum \text{all negative ranks } R_i = (-9) + (-7) + (-3) + (-8) + (-5) = |-32| = 32 \]

This observed test statistic value (\( P > .01 \)) is not small enough to disprove the null hypothesis (2 groups are equal) and thus it is accepted that there is not sufficient evidence to show that the two groups (UMMP family-level taxonomy vs. EcoAnalysts family level taxonomy) are different. Accepting the null hypothesis in this situation is critical since the identification of EPT families can be some of the most difficult. Also, the majority of UMMP sites are dominated by Trichoptera taxa and the Plecoptera and Ephemeroptera can be quite sparse in samples making their accurate family level identification essential to generating accurate biometrics for assessing river quality. Figure 2 graphically illustrates the EPT values generated from UMMP volunteer identifications compared with those families identified by professional entomologists at EcoAnalysts.

Figure 2: Percent EPT Comparison for UMMP vs. EcoAnalysts Data
The Family Biotic Index was the third and final biometric statistically analyzed with the Wilcoxon Signed-Rank Test for paired data. The Family Biotic Index takes into account the specific numerical tolerance values assigned to families within the aquatic insect community and generates a biotic index score for a particular site based upon the diversity and density of the benthic community present. Family tolerance values range from 0-10 with 0 being intolerant and 10 being the most tolerant. Precise family level identification of invertebrates is essential for calculating this biometric since the misidentification of an organism can mean the difference in recording a tolerance value of a 0 or a 10. This situation could completely shift the river quality score or assessment from one category to another (fair to good or good to excellent etc.).

Figure 3 presents the Family Biotic Index scores generated by the UMMP and EcoAnalysts data for the period covering 1995 through 1997. What is interesting to note about the results presented in Figure 3 is that on 8 out of the 11 comparisons presented in the graph, the EcoAnalysts generated biometrics for the Family Biotic Index scores were higher than those generated from the UMMP family level taxonomic data. Again, this reflects the larger density samples processed by EcoAnalysts, which inevitably will yield a higher number of family level taxa.

Figure 3: Comparison of Family Biotic Index Scores Generated with UMMP vs. EcoAnalysts Data

Although Figure 3 presents differences in Family Biotic Index scores graphically, when the scores are compared statistically with the Wilcoxon Signed-Rank Test for paired data, the differences have a degree of similarity strong enough to accept the null hypothesis that the two groups of Family Biotic Index scores are the same. Table 3 presents the UMMP and EcoAnalysts Family Biotic Index scores as well as the resulting statistical analyses from the signed rank test.
Table 3: Family Biotic Index

<table>
<thead>
<tr>
<th>U M M P</th>
<th>EcoAnalysts</th>
<th>Diff</th>
<th></th>
<th>Diff</th>
<th></th>
<th>Rank</th>
<th>Signed Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.7</td>
<td>4.3</td>
<td>-.6</td>
<td>.6</td>
<td>9.5</td>
<td>-9.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4.4</td>
<td>3.8</td>
<td>.6</td>
<td>.6</td>
<td>9.5</td>
<td>9.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4.2</td>
<td>4.5</td>
<td>-.3</td>
<td>.3</td>
<td>7</td>
<td>-7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4.9</td>
<td>5</td>
<td>-.1</td>
<td>1</td>
<td>2</td>
<td>-2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4.5</td>
<td>4.6</td>
<td>-.1</td>
<td>.1</td>
<td>2</td>
<td>-2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3.6</td>
<td>3.8</td>
<td>-.2</td>
<td>.2</td>
<td>4.5</td>
<td>-4.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3.6</td>
<td>3.5</td>
<td>.1</td>
<td>.1</td>
<td>2</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3.6</td>
<td>4.3</td>
<td>-.7</td>
<td>.7</td>
<td>11</td>
<td>-11</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3.6</td>
<td>3.9</td>
<td>-.3</td>
<td>.3</td>
<td>7</td>
<td>-7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5.4</td>
<td>5.7</td>
<td>-.3</td>
<td>.3</td>
<td>7</td>
<td>-7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5.5</td>
<td>5.3</td>
<td>.2</td>
<td>.2</td>
<td>4.5</td>
<td>4.5</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\[W = \left| W_+ \right| = \left| \sum_{all \ positive \ ranks} R_i \right| = 9.5 + 2 + 4.5 = 16\]
\[W = \left| W_- \right| = \left| \sum_{all \ negative \ ranks} R_i \right| = \left| (-9.5) + (-7) + (-2) + (-2) + (-4.5) + (-11) + (-7) + (-7) \right| = |-50| = 50\]

W is not small enough (where \(P > .10\)) to disprove the null hypothesis (H₀). Therefore, it is accepted that the two groups are the same.

The Wilcoxon Signed-Rank Test for paired data definitely proves that the biometric scores calculated from the UMMP and EcoAnalysts family level taxonomy are statistically similar in the case of the EPT and Family Biotic Index as Tables 2 and 3 clearly present. A less statistically driven assessment of the validity of the volunteer generated biometric scores can be determined simply by observing the difference columns in Tables 2 and 3 and noting that none of the differences calculated for these biometric scores exceeds 1. On a base level, this is an impressive statistic in itself given the high degree of variability encountered during the field and laboratory components of a volunteer biomonitoring program containing eleven stations. Not only does the statistical comparison validate the efforts of UMMP volunteers but it also provides a sound foundation for the data presented in the State of the Upper Merrimack 1995-1997.

Summary
A Wilcoxon Signed-Rank Test for paired data was applied to biometric scores that were calculated independently from family level invertebrate data compiled by the UMMP and EcoAnalysts respectively. Three biometrics were analyzed in this manner to determine whether or not there was a significant difference in the data generated from samples processed by UMMP volunteers versus those generated by professional entomologists. In two out of the three biometrics analyzed in this manner, the Wilcoxon Signed-Rank Test for paired data demonstrated that the biometric scores calculated by UMMP and EcoAnalysts were the same and not statistically different. The EPT and Family Biotic Index scores were determined to be statistical similar enough to not reject the null hypothesis that the biometric scores calculated with UMMP data and scores calculated with EcoAnalysts data are the same. Although the null hypothesis was rejected for the Family Richness biometric, this is not seen as a setback or as evidence for discounting the validity of the information presented in the State of the Upper Merrimack 1995-1997, a river quality report. On the contrary, this finding illustrated the fact that samples sent out to EcoAnalysts consistently contained a much higher organism density when compared to those analyzed by UMMP volunteers from the same site. This situation manifests from the fact that during the three hour “Bug-Night” workshops where volunteers process and conduct family level identifications, volunteers are often selecting samples that have the best
chance of being completed within the three-hour workshop interval. This resulted in most of the high-density samples being sent out to EcoAnalysts for identification and statistical analyses of three particular metrics.

This situation has provided the UMMP with an opportunity to adapt methodologies for validating volunteer generated data in the future. For example, in order to audit the family level identifications performed by UMMP volunteers, samples that have been completed by UMMP volunteers will be shipped out to an independent, professional laboratory for family level taxonomic verification. This will provide an audit in the purest sense instead of shipping out a replicate from the same site that has inherent differences in community structure to begin with. This is not to say that the information presented in this paper does not validate the data presented in the State of the Upper Merrimack 1995-1997. In fact, the statistical validity demonstrated from the biometric comparison calculated by the Wilcoxon Signed-Rank Test for paired data exemplifies the efforts and assessment information generated by a volunteer program that adheres strictly to an EPA approved Quality Assurance/Quality Control Plan. All reporting submitted by the UMMP should be considered by local and state agencies for inclusion in surface water quality assessments, local planning initiatives and as a vital component to developing and implementing a watershed approach for natural resource management and protection.

The results presented here not only validate the thousands of hours of volunteer monitoring and data generation associated with the UMMP but provide a sound statistical foundation for the recently published State of the Upper Merrimack 1995-1996 a river quality report. It is hoped that the information presented within this report, along with the results presented here will act as a template for other volunteer organizations to strive toward the goal of producing statistically valid data. The Upper Merrimack Monitoring Program will implement and improve the data auditing procedures that validate volunteer generated data as the program continues to monitor the physical, chemical and biological health of the upper Merrimack watershed into the 21st century.
References


Other Resources

Statistical analysis assistance from Stephen Eckberg.

E. coli samples processed by Franklin Waste Water Treatment Facility, Franklin, New Hampshire

Macroinvertebrate data processed with customized data processing templates created by River Watch Network, Montpelier, Vermont

Quality Control/Quality Assurance macroinvertebrate samples processed by EcoAnalysts, Moscow, Idaho

Random Quality Control/Quality Assurance E. coli samples processed by the NH Department of Environmental Services, Concord, New Hampshire

Topographic maps produced using TOPO! Software, Wildflower Productions, San Francisco, California

For further information about UMRLAC and UMMP, visit www.merrimackriver.org

Funded in part by a grant from the

NH Department of Environmental Services

Nonpoint Source Program Local Initiative Grant