Reducing the sensitivity of the water quality index to episodic events
Bruce W. Kilgour, Anthony P. Francis and Vincent Mercier

ABSTRACT
We demonstrated a general relationship between water quality index (WQI) values and indices of benthic community composition for a set of 32 streams from British Columbia and Ontario. Streams that produced lower WQI values tended to have benthic communities characteristic of degraded water quality. Streams, in contrast, that produced higher WQI scores tended to have a fauna characteristic of high water quality. Trimming the water quality data for high-total suspended solids (TSS) events increased the WQI values by as much as 30 points. There were modest but apparent increases in the strength of the association between the WQI and indices of benthic community composition when the water quality data records were trimmed of values that occurred during periods of high (extreme) turbidity. Trimming data that contained turbidity (or TSS) values beyond the mirrored 5th, 90th, or 95th percentile were about equal in their effect on the WQI. The removal of high TSS samples on the basis of the ‘mirror’ method can be recommended on the basis that it will likely correctly remove data that have a long right-hand tail, and will also correctly not remove data when the data are more normally distributed.

Key words | benthos, indicator, monitoring, suspended solids, water quality

INTRODUCTION
The Canadian Council of Ministers of the Environment (CCME) Water Quality Index (WQI) is a communication tool that summarizes water quality conditions at a site, much like an economic index (e.g., Standard & Poor’s Toronto Stock Exchange Index or S&P/TSX). The WQI is a representation of the number of parameters that exceed their guidelines, as well as the frequency and magnitude of those exceedances. Values range between 0 and 100, with higher values indicating water that tends to meet the guidelines more frequently, and that is considered to be of higher quality (Painter & Waltho 2003). The index is currently used by Environment Canada to summarize variations in water quality countrywide in our annual report on sustainability indicators (Environment Canada 2011). Additionally, many practitioners use the index to communicate the results of surface water studies to lay audiences. The index has been variously used across Canada (e.g., Khan et al. 2003, 2004, 2005; Lumb et al. 2006; Parparov & Hambright 2007; de Rosemond et al. 2009), as well as internationally (Rickwood & Carr 2009; Boyacioglu 2010).

Despite the communication benefits that the WQI provides, the current use of the index presents two issues. The first issue relates to interpreting the environmental or ecological consequences of index scores. Because the WQI is based on comparison of measured concentrations to water quality guidelines for the protection of aquatic life, the index is purported to be a measure of the quality of water and the inherent risks that surface water poses to aquatic organisms. There has, however, been no peer-reviewed or otherwise published accounting to demonstrate that the Canadian WQI is predictive of risks to aquatic organisms, or otherwise correlates to the ‘health’ (however defined) of aquatic organisms (but see Sylvestre & Ryan 2005, for a presentation at a conference). The first objective of this paper, therefore, is to document the degree of association between the WQI and the composition of benthic communities in...
riverine environments. Benthic invertebrates were selected for this aspect because they are an integral component of surface water ecosystems, because they are routinely used in monitoring programs all across Canada, they are known to reflect the quality of water (Lenat 1999) and because the condition of the benthic community can be used in general terms to predict the condition of other components (e.g., fish communities; Jackson & Harvey 1995; Kilgour & Barton 1999).

The second issue relates to the impact that extreme flow events have on WQI index scores. Concentrations of many water quality parameters in lotic waters (i.e., rivers, creeks) are related to flow volume and suspended-sediment concentration. Higher flow volumes cause erosion of bed sediments, and higher suspended sediment concentrations. High surface runoff from the watershed can also carry high sediment loads (Kerr 1995). Metals, nutrients (N and P), and coliforms are often associated with suspended sediments either because they are adsorbed to the surface of the sediments, or because they are simply associated with sediments in runoff from the landscape (Kerr 1995). Infrequent but high discharge events, therefore, can lead to high concentrations of metals, nutrients, and coliforms in samples collected from those time periods, and produce WQI values that imply poor water quality.

The CCME WQI is usually derived for riverine environments from a minimum of ~12 samples collected over a period of 3 years (i.e., four samples per year; Mercier & Léger 2005). The WQI does not account for samples collected from wet (precipitation days with high discharge) or dry (no precipitation) events, so the inclusion of a low-frequency high-discharge event as one of the 12 data records has the potential to cause a significant reduction in reported water quality with the index. It is, therefore, important to understand the consequences to the index of these rare events, and to determine if there are data treatments that can be used to minimize effects of these events on the index value. The second objective of this study, therefore, was to determine if there are data treatments that can be used to reduce the impact of extreme flow events on index values, and make the index more representative of the ‘quality’ of the water.

To achieve the two stated objectives, we sought data from monitoring programs that produced frequent water quality data, as well as collections of benthic macroinvertebrates. The British Columbia Ministry of the Environment, in coordination with Environment Canada, has monitored numerous rivers in British Columbia (BC), including the collection of frequent (every two weeks) water samples and annual collections of benthic macroinvertebrates at a number of locations. Toronto Region Conservation Authority and the Ontario Ministry of the Environment also collaborate in the Toronto region, collecting water samples at approximately monthly intervals in areas where benthic community samples are also collected annually. We worked with these agencies to identify appropriate site-level data, and to analyze those data to address the two stated objectives.

### METHODOLOGY

#### Available data

The Toronto and Region Conservation Authority (TRCA) provided water quality data from 23 locations (at which they also had collected benthic community data (Figure 1) in 2006). These locations are generally within urban stream systems. Water quality at these sites was characterized up to 12 times per year, without regard for whether the event being monitored was a ‘wet’ or ‘dry’ event. Measured water quality parameters for these sites included the following: Al, NH₃, Cl, Cu, Fe, Ni, TP, and Na. Benthic community samples were collected following the Ontario Ministry of Environment’s rapid bioassessment methodology described by Jones et al. (2007). Benthic samples were collected using D-framed kick nets (500-μm mesh), and comprised composites of kicks from two riffles and one pool. These TRCA data were used to explore the association between indices of benthic community composition and the WQI.

Environment Canada was able to provide data in BC for nine locations (Figure 1, data from 2005 and 2006). The BC locations are all rural interior sites. Water quality samples consisted of approximately bi-weekly sampling at each site, thus producing a relatively large data record (hundreds of records) for each sampled location. Benthic community samples were collected from each location following the
federal Canadian Aquatic Biomonitoring Network (CABIN) methodologies: basically timed (3-minute), travelling kick and sweep samples from riffle habitats using 400-μm mesh D-framed nets (Environment Canada 2012). The Environment Canada data were used to explore the associations between indices of benthic community composition and the WQI, and to explore the impact on the WQI of different data treatments for identifying potentially extreme flow events based on turbidity.

Computing indices of benthic community composition

In part because of the coarseness of the Ontario (ON) data, we computed simple indices of composition including percent of the fauna as EPT (i.e., mayflies, stoneflies, and caddisflies), as well as the percent of the fauna as each of other major groups (chironomids, oligochaetes, amphipods, etc.). Taxonomic indices were calculated using benthic species counts from one to three samples taken during the late summer to early fall in the final year of the defined WQI index period. Thus, benthic data were from either 2005 or 2006 for BC, or from 2006 for ON. The average count per taxon was used where more than one kick sample was available. Percent EPT is a common metric typically computed because those three groups are considered among the most sensitive to a variety of insults to water quality (Bode 1988). Chironomids and oligochaete worms, in contrast, are considered to be the groups most tolerant of degraded water quality (various causes; Bode 1988). The BC data were identified to lowest practical
levels. Taxa richness (number of taxa based on lowest taxonomic level as per CABIN, which is generally genus for common groups) was computed for the BC data. Taxa richness was not computed for the ON data because of its coarseness.

Computing the WQI

For the sites in BC and ON, we began by examining the list of parameters used to produce the existing WQI in each of the sites. The total list of water quality parameters used in any existing WQI was compiled to provide a list of parameters that, according to expert opinion, were important to water quality at least somewhere in BC. We then extracted from Environment Canada’s Pacific and Yukon Region Water Quality Monitoring Program website (http://waterquality.ec.gc.ca/EN/home.htm) records for as many of these parameters for each of our sites as possible from the years 2003 to 2006 inclusive. While data existed for most parameters at most sites, not all parameters were universally measured. Of the list of 28 parameters used in the calculation of WQI for at least one of our BC sites, only 13 parameters were commonly measured at each site (As, Cd, Cr, Cu, Fe, Pb, Mo, Ni, N, pH, TP, Th, Zn). In calculating WQI for each site, we used the CCME published guideline values for each parameter to define exceedances. However, as no common CCME guideline values exist for water temperature and phosphorus, temperature was also discarded as a contributor to our standard WQI and the Ontario Ministry of Environment guideline for total phosphorus (30 μg/L) was used (OMOE 1997). Thus, our standard WQI was based on 13 parameters. Two additional parameters were also collected, as they were important to subsequent WQI calculations and data treatments: hardness (CaCO₂) and turbidity.

The standard WQI (see equations below) was calculated for each site using water quality records for the year of the CABIN record(s) plus the previous 2 years. Thus, if CABIN data were available for 2006, the WQI was based on water quality data from 2004 to 2006 inclusive. The initial water quality data set consisted of records of all water samples for each site in the given 3-year period. Most records only included measurements on some parameters, often with multiple records existing per day, indicating replicates of the same measurement set and/or multiple samples covering different parameters. Site records were pooled over date, taking the mean where multiple measures of a parameter existed, to produce a single record per date per site. Finally, as hardness is a necessary covariable for determining guidelines for some parameters, data records lacking a measure of hardness were excluded. WQI was then calculated using the WQI calculator macro for each of the nine sites over the 3-year period using a total of 628 records. Not all of the records were complete. Some data records were missing a measurement on one or more parameters on a given day. Out of the total 628 records, 285 had at least one value missing. A similar procedure was carried out with data for sites in ON. We used the following water quality parameters as input into the WQI: Al, Cl, Cu, Fe, Ni, TP. A total of 367 data records were available for 27 sites (Figure 1).

We used the conventional calculation for the WQI provided by CCME (2001):

\[
CCMEWQI = 100 - \left( \frac{\sqrt{F_1^2 + F_2^2 + F_3^2}}{1.752} \right)
\]

(1)

\(F_1\) is the scope and represents the percentage of variables that do not meet their objectives at least once during the time period under consideration relative to the number of variables measured.

\[F_1 = \left( \frac{\text{Number of failed variables}}{\text{Total number of variables}} \right)
\]

(2)

\(F_2\) is the frequency and represents the percentage of individual tests that do not meet objectives (failed tests).

\[F_2 = \left( \frac{\text{Number of failed tests}}{\text{Total number of tests}} \right)
\]

(3)

\(F_3\) is the amplitude and represents the amount by which failed test values do not meet their objectives. There are three steps to its calculation, as follows:

\[F_3 = \left( \frac{nse}{0.01 \cdot nse + 0.01} \right)
\]

(4)
where, nse is the normalized sum of excursions, which is calculated as:

\[ nse = \frac{\sum_{i=1}^{n} \text{excursion}_i}{\# \text{ of tests}} \] (5)

where, an excursion is a measure of deviation from the water quality guideline. Excursions were computed as the following when the test value should not exceed the water quality guideline:

\[ \text{Excursion}_i = \frac{\text{Failed test value}_i}{\text{Objective}_i} \] (6)

Excursions were computed as the following when the test value should not be less than the water quality guideline:

\[ \text{Excursion}_i = \frac{\text{Objective}_i}{\text{Failed test value}_i} \] (7)

**Data trimming with turbidity**

The notion of filtering exceedances based on total suspended solids (TSS) can be considered valid for two reasons. First, high TSS can be expected to bind metals and other substances, making them less bioavailable, posing less risk to biological receptors. Second, in the event that substances are bioavailable during the infrequent high-discharge-related events, the briefness of the exposure (hours to a few days) is anticipated to pose a generally low risk to ecological receptors. This is particularly true considering that water quality guidelines are based on the most sensitive receptor at the most sensitive life stage, over an indefinite exposure period. Water quality data from BC and ON were ‘trimmed’ using three different criteria based on ‘extreme’ TSS events. Direct measures of TSS were available for ON samples but were highly limited within the BC database. Turbidity was used as proxy for TSS for BC. Turbidity generally has a strong linear relationship with TSS, although the slope of this relation will vary among sites (Hannouche et al. 2011). For this study, however, turbidity is used only to identify high TSS events relative to background TSS levels within streams individually, and thus coefficients of specific per stream TSS-turbidity relationships are not relevant. Each trimming method assumed that right-skews in the distribution of turbidity or TSS data reflected unusual flow events that occurred after a significant precipitation event, or during a spring melt event (Matschullat et al. 2000; Tri-Star 2006).

The first data-trimming method, here termed the ‘Mirrored 5th Percentile Method’, mirrors the upper portion of the data distribution to the lower portion of the data distribution. The method is derived from Matschullat et al. (2000), Tri-Star (2006), and Stantec Consulting and Coldwater Consulting (2008). The upper limit for the data is estimated as:

\[ \text{Upper limit} = \text{Median} + (\text{Median} - 5\text{th Percentile}) \] (8)

This method assumes that natural background data are distributed normally, and that the 5th percentile value is about as far to the left of the mode as the 95th percentile is to the right of the mode (Figure 2), and that this ‘estimate’ of the 95th percentile is a reasonable approximation to the limits of the normal range of variability. This method assumes that the lower tail of the data distribution were reasonable estimates of the lower tail of the distribution for natural background. In previous versions of this method (Matschullat et al. 2000; Tri-Star 2006; Stantec Consulting and Coldwater Consulting 2008), mirroring was done using the minimum value in the data set as the benchmark for the lower limits of the normal range. Minimum (or maximum) values are poorly estimated. We chose to use the 5th percentile because it is somewhat more precisely estimated (Berthouex & Hau 1991).

![Figure 2 | Conceptual illustration of the three data-trimming methods.](https://iwaponline.com/wqrj/article-pdf/48/1/1/379892/1.pdf)
The second data trimming involved removing data records that had turbidity or TSS values in excess of the 90th percentile of turbidity, while the third data trimming removed data records with turbidity or TSS in excess of the 95th percentile (Figure 2). We again chose to trim data sets that contained at least 12 data records. All of the sites from BC had hundreds of records and were thus trimmed. Nine of the 23 ON data sets met the 12-record criterion, and were thus trimmed, producing a total of 18 data sets for which we evaluated the effects of ‘trimming’ high TSS events.

**Correlation analysis**

We used simple inspection of biplots of data and bivariate correlation analysis to determine the nature and strength of the association between the WQI and relative abundances of major taxonomic groups (plus richness for the BC data). The ON and BC data were combined into one full data set to explore an overall association between the unadjusted WQI values and benthic taxa relative abundances. The decision to combine data was made after exploratory data analysis demonstrated that the two sets of data were complementary and assisted in understanding the underlying relationship between the two sets of variables (i.e., water quality and biology).

**RESULTS AND DISCUSSION**

Not all of the data records for any site were complete. That is, some data records were missing a measurement on one or more parameters on a given day. Out of the total 628 records, 285 had at least one value missing. This general lack of data consistency may cause WQI values to be artificially high. If an exceedance for one parameter, recorded on a given day, is accompanied by exceedances in other parameters, but those other parameters remain unmeasured due to lack of completeness of testing, potentially negative loadings on the WQI could be missed. The effect of missing data on the WQI, however, was not evaluated here.

Benthic communities from the Toronto area streams differed fundamentally from the benthos of the streams in BC. Toronto area streams had benthic communities dominated by chironomids and worms. The percent of the fauna as EPT taxa ranged between none and about 80% (i.e., a broad range), while there was only one site that contained stoneflies (Plecoptera). The streams in BC, in contrast, generally had lower relative abundances of chironomids (<30%) and worms (<25%) and variable percent of the fauna as EPT taxa (5–90%), and generally included a substantive number of stoneflies.

The BC sites produced a reasonable range of unadjusted WQI scores, varying between ∼48 and 90. Most of the BC sites had water with basic pH and moderate hardness. The

---

**Table 1** | Summary of water quality data for BC sites and the ‘untrimmed’ WQI values. Parameter values provided are mean concentrations

<table>
<thead>
<tr>
<th>Station name</th>
<th>Index period</th>
<th>Parameter values</th>
<th>Parameter</th>
<th>AS</th>
<th>Cd</th>
<th>Cu</th>
<th>Hard</th>
<th>Pb</th>
<th>Ni</th>
<th>pH</th>
<th>TP</th>
<th>Zn</th>
<th>WQI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Beaver</td>
<td>2004–2006</td>
<td></td>
<td>As μg/L</td>
<td>0.1</td>
<td>0.005</td>
<td>0.6</td>
<td>54</td>
<td>0.38</td>
<td>1.1</td>
<td>7.13</td>
<td>0.014</td>
<td>1.5</td>
<td>72</td>
</tr>
<tr>
<td>Fraser at Hope</td>
<td>2003–2005</td>
<td></td>
<td>Cd μg/L</td>
<td>0.8</td>
<td>0.050</td>
<td>3.6</td>
<td>57</td>
<td>0.73</td>
<td>4.6</td>
<td>7.78</td>
<td>0.074</td>
<td>6.4</td>
<td>48</td>
</tr>
<tr>
<td>Fraser at Red River</td>
<td>2004–2006</td>
<td></td>
<td>Cu μg/L</td>
<td>0.1</td>
<td>0.004</td>
<td>0.7</td>
<td>70</td>
<td>0.10</td>
<td>2.5</td>
<td>7.60</td>
<td>0.001</td>
<td>0.5</td>
<td>91</td>
</tr>
<tr>
<td>Kicking Horse</td>
<td>2004–2006</td>
<td></td>
<td>Hard mg/L</td>
<td>0.1</td>
<td>0.008</td>
<td>0.3</td>
<td>117</td>
<td>0.36</td>
<td>0.3</td>
<td>7.96</td>
<td>0.007</td>
<td>3.6</td>
<td>67</td>
</tr>
<tr>
<td>Kootenay</td>
<td>2004–2006</td>
<td></td>
<td>Pb μg/L</td>
<td>0.2</td>
<td>0.006</td>
<td>0.3</td>
<td>143</td>
<td>0.14</td>
<td>0.2</td>
<td>8.05</td>
<td>0.006</td>
<td>0.6</td>
<td>87</td>
</tr>
<tr>
<td>Quinsam</td>
<td>2004–2006</td>
<td></td>
<td>Ni μg/L</td>
<td>0.6</td>
<td>0.005</td>
<td>1.0</td>
<td>40</td>
<td>0.06</td>
<td>0.3</td>
<td>7.33</td>
<td>0.019</td>
<td>1.0</td>
<td>77</td>
</tr>
<tr>
<td>Salmon</td>
<td>2003–2005</td>
<td></td>
<td>Zn mg/L</td>
<td>1.0</td>
<td>0.034</td>
<td>1.8</td>
<td>165</td>
<td>0.42</td>
<td>1.5</td>
<td>8.07</td>
<td>0.079</td>
<td>3.2</td>
<td>64</td>
</tr>
<tr>
<td>Similkameen</td>
<td>2004–2006</td>
<td></td>
<td>TP mg/L</td>
<td>0.6</td>
<td>0.011</td>
<td>2.0</td>
<td>60</td>
<td>0.35</td>
<td>0.4</td>
<td>7.78</td>
<td>0.027</td>
<td>1.5</td>
<td>59</td>
</tr>
<tr>
<td>Sumas</td>
<td>2003–2005</td>
<td></td>
<td>Zn mg/L</td>
<td>1.0</td>
<td>0.018</td>
<td>2.2</td>
<td>135</td>
<td>0.23</td>
<td>37.7</td>
<td>7.77</td>
<td>0.097</td>
<td>2.7</td>
<td>48</td>
</tr>
<tr>
<td>CCME guideline</td>
<td></td>
<td></td>
<td>WQI</td>
<td>5</td>
<td>0.017</td>
<td>2–4</td>
<td>1–7</td>
<td>25–150</td>
<td>6.5–9.5</td>
<td>0.025</td>
<td>30</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*aDepends on hardness.*
Sumas River had the most degraded water quality, producing a WQI of 48, caused principally by high total phosphorus concentrations. Other sites with high total phosphorus concentrations included the Salmon River and Similkameen River (Table 1). Other parameters were within guidelines when water hardness was taken into account. Water quality data from the Toronto area streams produced generally lower WQI values ranging between 22 and 68. Toronto area streams had high total phosphorus concentrations as well as high aluminum, chloride, copper, and iron (Table 2). Other parameters were measured, but not reported here.

The WQI was correlated to a variety of indices of benthic community composition (Table 3). Benthic indices typically associated with water of higher quality tended to be more prevalent in streams that produced higher WQI scores. The percentage of the fauna as EPT (mayflies, stoneflies, and caddisflies), for example, was generally a larger fraction of the

### Table 2: Summary of water quality data for ON sites and the 'untrimmed' WQI values. Parameter values provided are mean concentrations

<table>
<thead>
<tr>
<th>Station name</th>
<th>Index period</th>
<th>parameter</th>
<th>Al μg/L</th>
<th>Cl mg/L</th>
<th>Cu μg/L</th>
<th>Fe μg/L</th>
<th>Ni μg/L</th>
<th>TP mg/L</th>
<th>WQI</th>
</tr>
</thead>
<tbody>
<tr>
<td>06008000202</td>
<td>2004–2006</td>
<td></td>
<td>394</td>
<td>317</td>
<td>4</td>
<td>626</td>
<td>3</td>
<td>0.120</td>
<td>36.0</td>
</tr>
<tr>
<td>06008000602</td>
<td>2004–2006</td>
<td></td>
<td>308</td>
<td>968</td>
<td>5</td>
<td>502</td>
<td>3</td>
<td>0.089</td>
<td>33.1</td>
</tr>
<tr>
<td>06008200302</td>
<td>2004–2006</td>
<td></td>
<td>574</td>
<td>1,523</td>
<td>6</td>
<td>780</td>
<td>5</td>
<td>0.115</td>
<td>31.7</td>
</tr>
<tr>
<td>06008300902</td>
<td>2004–2006</td>
<td></td>
<td>67</td>
<td>36</td>
<td>3</td>
<td>305</td>
<td>2</td>
<td>0.059</td>
<td>28.3</td>
</tr>
<tr>
<td>06008301902</td>
<td>2004–2006</td>
<td></td>
<td>337</td>
<td>281</td>
<td>4</td>
<td>248</td>
<td>2</td>
<td>0.057</td>
<td>57.4</td>
</tr>
<tr>
<td>06008310302</td>
<td>2004–2006</td>
<td></td>
<td>1,301</td>
<td>144</td>
<td>4</td>
<td>1,399</td>
<td>4</td>
<td>0.118</td>
<td>40.6</td>
</tr>
<tr>
<td>06008310402</td>
<td>2004–2006</td>
<td></td>
<td>46</td>
<td>47</td>
<td>2</td>
<td>172</td>
<td>2</td>
<td>0.062</td>
<td>24.8</td>
</tr>
<tr>
<td>06008500402</td>
<td>2004–2006</td>
<td></td>
<td>632</td>
<td>603</td>
<td>6</td>
<td>921</td>
<td>6</td>
<td>0.363</td>
<td>66.9</td>
</tr>
<tr>
<td>06008501402</td>
<td>2004–2006</td>
<td></td>
<td>56</td>
<td>574</td>
<td>11</td>
<td>348</td>
<td>4</td>
<td>0.135</td>
<td>24.4</td>
</tr>
<tr>
<td>06009400202</td>
<td>2004–2006</td>
<td></td>
<td>125</td>
<td>454</td>
<td>4</td>
<td>332</td>
<td>6</td>
<td>0.047</td>
<td>31.5</td>
</tr>
<tr>
<td>06009701102</td>
<td>2004–2006</td>
<td></td>
<td>291</td>
<td>285</td>
<td>3</td>
<td>362</td>
<td>5</td>
<td>0.059</td>
<td>41.2</td>
</tr>
<tr>
<td>06009701302</td>
<td>2004–2006</td>
<td></td>
<td>155</td>
<td>104</td>
<td>5</td>
<td>230</td>
<td>6</td>
<td>0.038</td>
<td>48.8</td>
</tr>
<tr>
<td>0610400102</td>
<td>2004–2006</td>
<td></td>
<td>367</td>
<td>54</td>
<td>2</td>
<td>457</td>
<td>4</td>
<td>0.068</td>
<td>47.3</td>
</tr>
<tr>
<td>0610400802</td>
<td>2004–2006</td>
<td></td>
<td>209</td>
<td>47</td>
<td>2</td>
<td>384</td>
<td>4</td>
<td>0.083</td>
<td>51.5</td>
</tr>
<tr>
<td>06010700202</td>
<td>2004–2006</td>
<td></td>
<td>671</td>
<td>68</td>
<td>3</td>
<td>733</td>
<td>9</td>
<td>0.058</td>
<td>59.7</td>
</tr>
<tr>
<td>8th Concession</td>
<td>2004–2006</td>
<td></td>
<td>96</td>
<td>117</td>
<td>3</td>
<td>158</td>
<td>4</td>
<td>0.050</td>
<td>43.6</td>
</tr>
<tr>
<td>97777</td>
<td>2004–2006</td>
<td></td>
<td>342</td>
<td>324</td>
<td>3</td>
<td>575</td>
<td>4</td>
<td>0.069</td>
<td>47.7</td>
</tr>
<tr>
<td>Brock Ridge</td>
<td>2004–2006</td>
<td></td>
<td>214</td>
<td>53</td>
<td>2</td>
<td>321</td>
<td>4</td>
<td>0.052</td>
<td>55.6</td>
</tr>
<tr>
<td>DM 6.0</td>
<td>2004–2006</td>
<td></td>
<td>372</td>
<td>606</td>
<td>7</td>
<td>580</td>
<td>6</td>
<td>0.095</td>
<td>31.1</td>
</tr>
<tr>
<td>DN008WM</td>
<td>2004–2006</td>
<td></td>
<td>1,425</td>
<td>206</td>
<td>8</td>
<td>1,947</td>
<td>10</td>
<td>0.056</td>
<td>25.9</td>
</tr>
<tr>
<td>E 2.4</td>
<td>2004–2006</td>
<td></td>
<td>8</td>
<td>939</td>
<td>7</td>
<td>154</td>
<td>1</td>
<td>0.052</td>
<td>42.8</td>
</tr>
<tr>
<td>HU010WM</td>
<td>2004–2006</td>
<td></td>
<td>746</td>
<td>110</td>
<td>4</td>
<td>938</td>
<td>8</td>
<td>0.055</td>
<td>33.5</td>
</tr>
<tr>
<td>Mayfield</td>
<td>2004–2006</td>
<td></td>
<td>199</td>
<td>99</td>
<td>4</td>
<td>415</td>
<td>20</td>
<td>0.088</td>
<td>33.7</td>
</tr>
<tr>
<td>MM003WM</td>
<td>2004–2006</td>
<td></td>
<td>836</td>
<td>412</td>
<td>8</td>
<td>1,168</td>
<td>9</td>
<td>0.068</td>
<td>24.8</td>
</tr>
<tr>
<td>RG007WM</td>
<td>2004–2006</td>
<td></td>
<td>440</td>
<td>79</td>
<td>2</td>
<td>483</td>
<td>8</td>
<td>0.040</td>
<td>44.1</td>
</tr>
<tr>
<td>RG008WM</td>
<td>2004–2006</td>
<td></td>
<td>591</td>
<td>131</td>
<td>4</td>
<td>658</td>
<td>8</td>
<td>0.070</td>
<td>34.1</td>
</tr>
<tr>
<td>Shoal Point</td>
<td>2004–2006</td>
<td></td>
<td>35</td>
<td>112</td>
<td>4</td>
<td>1,152</td>
<td>2</td>
<td>0.057</td>
<td>43.9</td>
</tr>
<tr>
<td>CCME Guideline</td>
<td>5–100^a</td>
<td></td>
<td>240</td>
<td>2–4^a</td>
<td>300</td>
<td>25–150^a</td>
<td>0.025</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

^a Depends on pH or hardness.
benthic fauna in streams that produced a higher WQI. Similarly, stoneflies were present only in streams that produced a WQI score of 60 or better. The percentage of the fauna as Diptera (a category of miscellaneous larval flies) increased with the WQI similarly to percent EPT. Miscellaneous Diptera include taxa like Tipulidae (crane flies) and Tabanidae (horse flies), groups that are relatively sensitive to perturbation and degradation of water quality (Bode 1988; Mandaville 2001). Other groups that are more tolerant of degraded water and habitat quality, such as oligochaete worms and chironomids, comprised a larger fraction of the fauna in streams that had lower WQI scores. These findings are consistent with the expectations of the WQI.

The strength of the relationships between the WQI and indices of benthic community composition were somewhat weak. Correlation coefficients (r-values) were generally <0.5, indicating that variations in the WQI explained about 25% of the variation in some of the indices of benthic community composition. Correlation coefficients would have been higher had there been a broader range of WQI values (i.e., lengthening the x-axis forces a stronger relationship when one exists). Notable in the biplots for chironomids, worms (Oligochaeta) and flies (Diptera) were generally ‘wedge-shaped’ distributions (Figure 5). Relationships between a biological receptor and a physical/chemical predictor like the WQI that produce a wedge-shaped distribution indicate the presence of other unmeasured variables that have not been accounted for or controlled (Cade & Richards 2005). The percent of the fauna as oligochaete worms, for example, was apparently controlled by water quality when the quality of the water was high (i.e., high WQI values), and less controlled by water quality when WQI values were low. At low WQI values, other habitat-related factors appear to have imposed some control on numbers of worms. There was no obvious influence of substrate on the composition of benthic fauna in this analysis, although there were only three sites that had a substrate dominated by fine material (Figure 3). We did not examine a possible influence of season on benthic community composition, but, with indices computed using ‘major-group’ taxonomy, the influence of season could be expected to be minor in relation to water quality and substrate influences (Borisko et al. 2008). Substrate, otherwise, is known to be a principal factor influencing the kinds of benthos found in a stream sample. Other potentially important modifiers included flow velocities and river size (e.g., bankfull width), but the data were not available to examine those influences, while they would have a secondary influence to substrate texture and water quality. The fact that relationships between the WQI and composition of the benthic community were observed in this analysis indicates that general conclusions of condition of benthos in a river can be inferred from computed estimates of the WQI.

All three statistical criteria for trimming events with high TSS from the data record produced a modest increase in WQI scores (Figure 4), with the magnitude of the increase depending somewhat on the data set. The WQI values for sites in BC increased in value between 10 and 30 points, whereas the increases in WQI values for streams in Toronto were only about 20 points (maximum). Trimming may have a larger impact when a stream has high water quality (i.e., generally has few exceedances of water quality guidelines) in low-discharge periods. The streams in Toronto had benthic communities indicative of degraded conditions, as well as lower WQI values suggesting degraded conditions, regardless of trimming extreme flow events.
The statistically significant correlations for the original WQI involved taxa richness, % EPT, % Plecoptera, and % Diptera. The strength of the correlation between the WQI and indices of composition increased in almost all cases when data trimming was used (Table 3). All data-trimmed WQIs were about equally correlated with the various indices of benthic community composition. A much larger set of locations would be needed to confirm with enough statistical power which of the data-trimming methods was superior. These findings, however, can cautiously be interpreted to suggest that the reduction in WQI values via the data filtering process does result in WQI values that are more predictive of the condition of the benthic community.
Changes in the WQI with trimming were associated with changes in the F1 and F3 components of the index. Figure 5 illustrates the changes in the three F components. The largest changes occurred in F1 (fraction of variables failing) and F3 (magnitude of excursions), with negligible changes in the F2 component (fraction of tests failing).

The negative relationship between taxa richness and the WQI is not overly unusual, but does somewhat contradict the notion that an improvement in water quality (i.e., increase in WQI value) should also result in an increase in the number of taxa. Relations between benthic indices such as abundance and richness are often modal (i.e., there is an optimum) in relation to some stressor. Pearson & Rosenberg (1978), for example, demonstrated that modest increase in nutrient enrichment marine offshore environments can lead to an increase in abundance, richness, and biomass of marine benthic macroinvertebrates. That same phenomenon was also demonstrated in freshwaters. Lowell et al. (2005) showed that modest enrichment in fresh surface waters, caused by pulp and paper mill effluents, can lead to an increase in abundance and richness of freshwater benthic macroinvertebrates.

The observation of stronger correlations with data-trimmed WQIs suggests a conclusion that concentrations of metals and other substances during periods of high discharge (i.e., extreme TSS events) do not pose a significant risk to benthic macroinvertebrates. We can only speculate that the risk is lower because substances are in a bound form with suspended particulate matter or because the exposure period is too brief. Regardless, this observation strongly suggests that WQI values should be computed with extreme flow events trimmed from the data set.

Here, we used turbidity to trim the BC data and total suspended particulate matter to trim the ON data. There were larger changes in the BC WQI values than in the ON WQI values. Again, and as above, the more subtle changes in the ON WQI values may have been a result of the ON sites having generally poorer water quality. We can only speculate with these data that trimming the data with turbidity or TSS will produce an equivalent

Figure 5 | Biplots of raw and adjusted F factors.
result on the WQI, as surrogate indicators of extreme flow events.

All of the methods described here for trimming high TSS events increased the WQI and potentially improved the ability of the WQI to predict the condition of a benthic invertebrate community. The potentially most applicable of the three methods is the mirror method that trims values that exceed the median plus the difference between the median and 5th percentile value. Most water quality concentration data are skewed, with long tails on the right-hand side. The mirror method will effectively trim data that are skewed in this way. In unusual circumstances, water quality data may be skewed with long tails to the left, and it would be the data to the left that would be considered unusual. If this were the case, the mirror method would appropriately not trim any data, while the two methods based on extreme right-hand percentiles (95th and 90th) would trim data (5 or 10%, respectively). On this basis, then, we propose that the mirror method be considered for routine identification and removal of high TSS water samples.

Data trimming may be problematic in some instances. If, for example, high TSS conditions (unrelated to precipitation events) were to occur for an extended period, say weeks to months, and if we had the luxury of an extended data set that had frequent data collected over several years, the removal of the high-TSS event data could lead to an underestimate of ecological risk. It is, therefore, important to review the time trend of TSS data to determine if removal of those events could potentially strip a relatively long event that could pose potential ecological risks.

An alternative means of ‘trimming’ data not considered here is based on precipitation events. It is relatively routine practice in ON to classify water quality data into ‘wet’ and ‘dry’ events (e.g., Booty et al. 2005). Wet event samples with high TSS typically produce ‘poorer’ WQI values, while dry event samples with low TSS typically produce ‘better’ (higher) WQI values. Thus, classification of sites using precipitation data would also be effective. The use of wet and dry events to classify data would assist in identifying those highly TSS events that are not precipitation related.

Some consideration for data transformations is required when the mirror method is applied to the data. Log-normal distributions (i.e., long right-hand tail) are frequently encountered in water quality data. Log-transformation of those kinds of data typically results in a ‘normal’ distribution of the concentrations. A mirror method applied to log-transformed data will, de facto, not identify as many extreme high-TSS events as being unusual. In the application of the mirror method here, the data were not log transformed. On the basis of the experience of this analysis then, we recommend that the mirror method be applied to non-transformed data.

CONCLUSIONS

This study has demonstrated a relatively robust relationship between the CCME WQI and indices of benthic community composition in stream environments. High WQI scores coincided with benthic communities that were diverse and dominated by relatively sensitive organisms. Lower WQI scores co-occurred with benthic communities that were less diverse and dominated by more tolerant forms such as worms and chironomids. Stripping of the data potentially improves the strength of the relationship between the WQI and indices of benthic community composition, while the three data-stripping methods were about equal in the improvement they provided. A much larger data set would be required in order to identify the optimal method. The removal of high TSS samples on the basis of the ‘mirror’ method can be recommended on the basis that it will likely correctly remove data that have a long right-hand tail, and will also correctly not remove data when the data are more normally distributed.

ACKNOWLEDGEMENTS

This work was funded by the Canadian Council of Ministers of the Environment (CCME) under contract to Kilgour & Associates Ltd. Thanks to L. Hamilton for help with revisions to the manuscript.

REFERENCES


Tri-Star 2006 The Statistical Approaches and Data Availability for a Case Study Comparison – Natural Background Levels and the CCME WQI. Report to Environment Canada by Tri-Star Environmental Consulting.

First received 7 February 2012; accepted in revised form 19 September 2012.