5.2.2 Summary Description

The evaluation of the macroinvertebrate data by station and by reach found similar results; the macroinvertebrate community is dominated by a few opportunistic Diptera (Chironomidae) and non-insect taxa (Oligochaetes). Nearly half of the taxa collected in the CAWS are from the order Diptera, and almost all are in the family Chironomidae. By abundance, oligochaetes (Phylum Annelida) dominate the benthic community, comprising over 74 percent of all macroinvertebrates collected from the CAWS during the 2001-2007 period. Two species of non-native bivalve, the zebra mussel, *Dreissena polymorpha*, and the closely related Quagga mussel, *Dreissena rostriformis bugensis* comprise 15 percent of the samples as well.

An analysis of the differences between sampling methods, i.e. grab samples (ponar) and artificial substrate samples (Hester-Dendy), show that richness measures (total richness, EPT richness, and diptera richness) are higher in the Hester-Dendy samples. In contrast, EPT taxa were nearly absent from the ponar collections with EPT richness values of zero for most ponar samples showing that the two sampling methods collected different organisms and in different quantities. The lack of EPT taxa in ponar samples suggests that lack of suitable substrate is a physical habitat limitation for benthic invertebrates. The presence of intolerant benthic EPT taxa in Hester-Dendy samples and the absence of EPT taxa in Ponar samples suggest sediment toxicity to mayfly, stonefly, and caddisfly larvae.

An analysis of macroinvertebrate metrics appropriate for evaluation within the CAWS was conducted. This analysis included a correlation analysis of macroinvertebrate metrics with sediment contamination. Five metrics were identified based on their sensitivity to contaminated sediments. These are taxa richness, percent Diptera, percent Oligochaetes, percent shredders and function feeding group diversity. The CAWS contains legacy contaminants that likely influence the metrics. The Hester-Dendy technique is sampling a population that is less exposed to environmental stress than the ponar sampling technique, which samples invertebrate communities in direct contact with sediments. The community differences were identified by a comparative analysis of the two sampling methods, which varied by metric and monitoring station. For example, ponar sampling resulted in lower species richness dominated by pollution tolerant individuals (oligocheates).

Additionally, an analysis of the macroinvertebrate dataset of the percent of head capsule deformities of larvae of the Chironomidae family (midges) was conducted within the Study Area for the 2001-2007 period. Deformities in midge larvae head capsules have been frequently observed in contaminated sediments. Deformity is generally considered to be a sublethal, teratogenic response to contamination. In an analysis of variance test, we concluded that there is no significant difference between mean rates of head capsule deformities for those collected on Hester-Dendy samplers and those collected in ponar dredge samples ($F=2.89$, $p=0.0911$). The strengths of correlation were significant ($p<0.05$) in the Hester-Dendy samples for ammonia-N (r=-0.399), iron (r=0.361), and DDx (DDT + DDE + DDD) (r=-0.396). Spearman correlation coefficients were significant for the ponar samples for mercury (r=0.659), cadmium (r=0.339), copper (r=0.439), simultaneously extracted metals (SEM) (r=0.455), SEM-acid volatile sulfides (r=0.454), total PCB (r=0.316) and semi-volatile organic compounds (r=0.323). No contaminants displayed strong correlations for both collection methods. This may reflect differences in exposure routes or pathways for macroinvertebrates in ponar samples and Hester-Dendy samples.

6. HABITAT DATA ANALYSIS

As discussed in Section 2.5, the process used to analyze habitat data in the CAWS and to develop a CAWS-specific habitat index was based on the process used to develop a non-wadeable habitat index (NWHI) for Michigan (Wilhelm et al., 2005). The process involves three major elements:

1. Sequential reduction of the list of habitat variables using qualitative screening, correlation analysis, and principle components analysis;

2. Identification of the key habitat variables that best explain fish data using multiple linear regression; and.

3. Incorporation of the key habitat variables into an index that can be applied to measure variation and change in the system.

This section describes the processing and analysis of habitat data for these purposes.

6.1 IDENTIFICATION AND SCREENING OF HABITAT VARIABLES

Based on review of the Wilhelm paper (Wilhelm et al., 2005); other relevant technical literature (Arlinghaus et al., 2002; Wolter and Arlinghaus, 2003; Short et al., 2005; Tate et al., 2005), data collected by the District as part of the ambient water quality monitoring program, and firsthand observations of conditions in the CAWS, a list of 242 habitat variables was compiled as a starting point. The starting list of 241 habitat variables is presented in Appendix E and is organized into five categories: geomorphology and hydrology; sediment and substrate; in-stream and riparian cover; bank and riparian condition; and anthropogenic factors.

Because the ultimate objective was to use multiple linear regression to analyze the CAWS habitat data with CAWS fish data, it was necessary to reduce the number of habitat variables substantially. Using the District data from 2001 through 2007, there were 81 paired sets of habitat and fish data. Multivariate statistical analyses require that the ratio of variables to data be as low as possible. It has been suggested that, for analysis of ecological data, the variable-to-data ratio be 0.1, but may be as high as 0.5 (Smogor and Angermeier, 1999). This rule of thumb suggests that the number of habitat variables in this Study should be reduced to somewhere between 8 and 40, preferably closer to the low end of this range to yield a ratio close to 0.1. The stepwise process used to reduce the list of habitat variables to a suitable number for multiple linear regression is described in Figure 6-1 and described in detail in Appendix D.

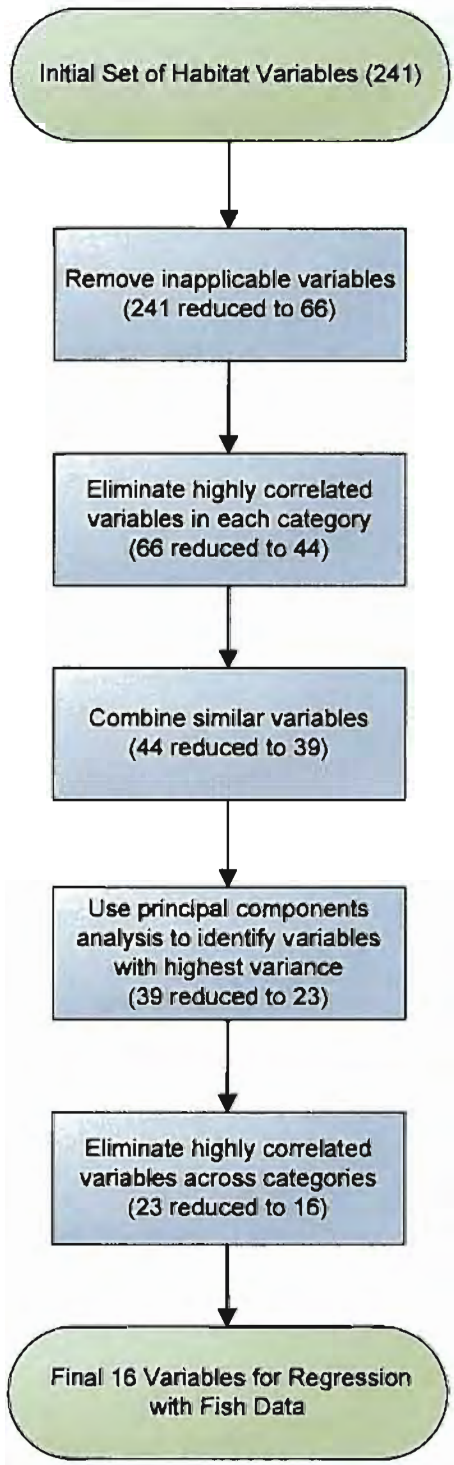

Figure 6-1: Process Used to Reduce the Set of Habitat Variables for Analysis with Fish Data.

This process outlined in Figure 6-1 was effective in reducing the set of habitat variables to 16, which represented a variable-to-data ratio of about 0.2.

Table 6-1: Final Set of Habitat Variables for Regression with Fish Data.

Variable Category	Habitat Variable
Geomorphology & Hydrology	Flashiness Index Wetted perimeter of channel Maximum depth in reach Number of off-channel bays Bank "pocket" areas
Sediment & Substrate	% Gravel, cobbles, boulders, shallow % Gravel, cobbles, boulders, deep % Plant debris on bed % Organic sludge
In-Stream Cover	Average macrophyte cover In-stream cover present Secchi depth
Bank & Riparian Condition	Dominant riparian land use % Vertical walled banks in reach % Riprap banks in reach
Anthropogenic Impacts	Manmade structures

These 16 variables were carried forward for comparison to fish data, described below.

6.2 ANALYSIS OF THE RELATIONSHIP BETWEEN FISH AND PHYSICAL HABITAT IN THE CAWS

The process described in Section 6.1 and Appendix D effectively reduced 241 potential habitat variables to a much smaller set of 16, that represented the habitat variables with the least inter-variable correlation and which explained most of the variance in the habitat data set. The next task in this analysis was to analyze the relationship of these variables to fish in the CAWS. There were several objectives for this, including the following:

- Determine which physical habitat variables are the most significant to fish in the CAWS.
- Determine how much of the variability in the CAWS fish data can be explained by physical habitat.
- Compare the relative importance of physical habitat to fish in the CAWS, with that of water quality.

Statistical analysis of the fish and habitat data from the CAWS was used to attain these objectives. Specifically, multiple linear regression was used to compare habitat variables to paired fish data to determine which of the 16 habitat variables best explain variability in fish data in the CAWS. The methodology and results of this analysis are described below.

6.2.1 Methodology

Various methods can be used for comparing fish data and habitat data from a single system to address the objectives listed above. Review of the professional literature related to assessment of aquatic habitat shows a range of dependent variables and mathematical methods have been used and published in the peer-reviewed literature. No commonly accepted standards have been developed for this type of analysis, so selection of the methodology must rely to a large extent on professional judgment. In this study, the methods selected were based on the needs of the study, review of methods used by other investigators in similar studies, and on understanding of the unique aspects of the CAWS. More details on the methodology used are presented below.

6.2.1.a Representation of Fish Data in the Analysis of Habitat Data

As discussed in Section 2.5, fish were selected as the indicator biota for comparison to physical habitat data in this Study. Twelve key fish metrics were identified (Appendix A) using CAWS fish data collected by the District between 2001 and 2007 (Table 6-2). For purposes of comparing these fish metrics to habitat data, it was necessary to combine the metrics into a single value. A fish index of biological integrity (IBI) was not available that incorporated the selected metrics, although the process used to select the fish metrics was exactly the same process used in many fish IBI studies.

Statistical comparison of habitat variables with each of the twelve fish metrics would have been cumbersome and might not have yielded conclusive results regarding which habitat variables were most important to understanding fish data in the CAWS. So, as a starting point, the fish metrics were divided into the five ecological function categories and compared to habitat variables using multiple linear regression. Each of the fish metrics was first transformed to a normal distribution, if necessary, and standardized to give each metric equal weight. Then the metrics within each functional category were simply summed. Metrics that reflected positive conditions were assigned a positive value and metrics that reflected a negative condition were assigned a negative value.

Table 6-2: Selected CAWS Fish Metrics.

Fish Metric	Ecological Function Category
% Diseased or with eroded fins, lesions, or tumors	Abundance and condition metric (ACM)
catch per unit effort	Abundance and condition metric (ACM)
% lithophilic spawners by count	Reproductive function metric (RFM)
% insectivores by count	Trophic function metric (TFM)
% top carnivores by weight	Trophic function metric (TFM)
proportion of Illinois tolerant species	Indicator species metric (ISM)
IL ratio of non tolerant coarse-substrate spawners	Reproductive function metric (RFM)
number of IL native minnow species	Species richness and composition metric (SRC)
number of IL native sunfish species	Species richness and composition metric (SRC)
IL ratio of generalist feeders	Trophic function metric (TFM)
% Intolerant species by count	Indicator species metric (ISM)
% moderately intolerant species by weight	Indicator species metric (ISM)

This process showed that, when grouped by function, the ACM metrics (catch per unit effort and percent diseased or with eroded fins, lesions, or tumors) had relatively weak correlation with habitat. The other four functional categories were approximately equal in their relationship to habitat. Based on these observations, a combined fish metric was calculated by summing the reproductive function, trophic function, indicator species, and species richness and condition metrics. Because a system-specific index of biotic integrity (IBI) for fish does not exist for the CAWS and other IBIs are not appropriate for the CAWS (see Appendix A) this combined fish metric was used in subsequent analyses with habitat data.

6.2.1.b Determination of Habitat Variables for Study Period

It would not be feasible to conduct this Study at present without relying on the data collected by the District in the past, as these data provide valuable measures of CAWS fisheries over many years. However, only a relatively limited set of physical habitat data were measured concurrent with the District's fish sampling events from 2001 through 2007. Therefore, to use the District's fish data in this Study, it was necessary to make some assumptions regarding physical habitat during that time period, as described below.

- All hydrologic variables were assumed constant from year to year, using model predictions the DUFLOW model developed by Marquette University. Given the highly regulated hydrology of the system and the fact that most of the flow entering the CAWS is from wastewater treatment plants, it is unlikely that significant variations in average or extreme hydrologic variables occurs from year to year.

- Bank and riparian conditions were assumed to be the same as observed in 2008, unless otherwise noted in the District's physical habitat observations. Given the urban, constructed nature of the CAWS, this is likely a safe assumption. No major changes in these conditions were noted in consultation with District personnel involved in routine monitoring in the CAWS.

- No quantitative measurements of macrophyte growth were available from 2001 – 2007. Quantitative measurements of littoral macrophyte coverage were made in 2008 as part of this Study, though, and the presence of aquatic macrophytes was noted on the historical habitat assessment forms completed by the District from 2002-2007. Lacking historical data, but recognizing the probable importance of macrophyte cover, the decision was made to retroactively apply 2008 macrophyte measurements to the period of 2001 – 2007. While this is likely not an accurate representation of historical conditions, it is better than disregarding macrophytes altogether. Furthermore, review of the historical habitat assessment forms generally corroborated the 2008 data.

In this Study, the assumptions regarding the similarity of physical habitat condition between 2008 and the preceding seven years are believed to be reasonable, given the relatively unchanging nature of conditions within the CAWS and the nature of the subject variables. The percentage of vertical walled banks at a sampling station, for example, was likely about the same in 2008 as it was in 2001. Although minor changes cannot be ruled out, they are likely not significant compared to the variability in fish data at these stations from year to year, which can be quite large.

One variable that is less reliably estimated in this retroactive manner is Secchi depth, which was not measured during 2001 – 2007, but was measured in 2008 for this study. Historical turbidity data collected by the District shows that water clarity can vary over time in the system, so assuming that 2008 Secchi measurements accurately reflect conditions at a location in preceding years is probably not accurate. As an alternative, 2008 Secchi data were compared to turbidity measurements from the CAWS to assess whether historical Secchi could be estimated using turbidity (Figure 6-2).

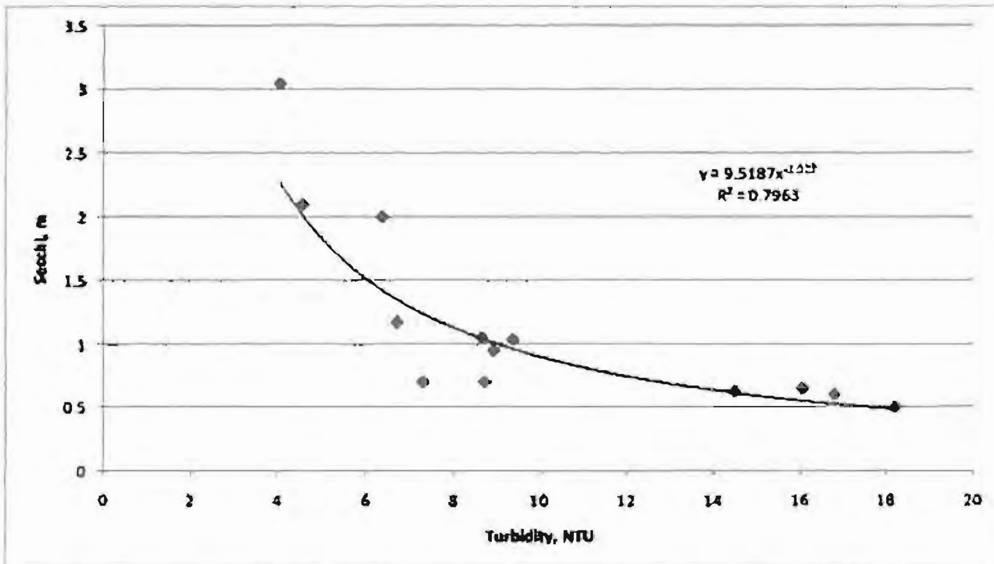

Figure 6-2: Comparison of 2008 Secchi Measurements with 2008 Turbidity Measurements.

The regression of the 2008 Secchi with the 2008 turbidity yields an r-squared value of nearly 0.8, which indicates a relatively strong relationship between the two measurements. However, there is still as much as a 0.5 m variance between actual and predicted Secchi using the regression relationship, which could result in a prediction error of approximately 50% for areas where Secchi is on the order of 1 meter depth, which is common in the CAWS. In addition, Secchi is typically used in habitat studies as an indicator of light penetration, related to the growth of aquatic macrophytes that create fish habitat and provide food. In this Study, a metric reflecting macrophyte growth was already included, so Secchi was, in this sense, redundant. For these reasons, Secchi was eliminated from the analysis, which resulted in 15 habitat variables for the regression analysis.

6.2.1.c Description of Multiple Linear Regression Method Used

For this analysis, multiple linear regression (MLR) was chosen as the statistical method for comparing habitat variables with fish data, for a number of reasons. First, MLR is a mathematically rigorous method that has been used in several habitat studies published in professional literature and for development of habitat indices. Second, MLR was used in the development of the Michigan Non-Wadeable Habitat Index, which was the model approach for this study as discussed in Section 2.5 of this report. Third, MLR provides a parametric measure of goodness-of-fit (i.e., r-squared value) that allows relatively straightforward comparison of data models to each other and that provide a quantitative measure of the degree to which the independent variable data (i.e., habitat or water quality) describe the variation in the dependent variable data (i.e., fish data).

Several MLR methods exist to choose from. The most commonly used methods are standard stepwise, forward selection stepwise, backward elimination stepwise, and best subsets. Each of the three stepwise methods involves starting with an initial set of variables in the regression model and then adding or removing variables according to a set of rules until some subsequent steps do not improve the fit of the model to the data. The best subsets method calculates all possible regression models using all possible numbers of variables. Instead of producing a single regression model, the best subsets method produces several to choose from.

Stepwise regression methods have been criticized because they do not allow the application of specialized knowledge about the data or the system being studied to inform the selection of the regression model. For this reason, the best subsets method was selected for this study. As will be shown in subsequent sections of this report, this method produced several possible regression models that allowed the opportunity for comparison between models and the application of judgment regarding model selection.

6.3 SYSTEM-WIDE COMPARISON OF HABITAT WITH FISH

The final selected set of habitat variables were compared to the CAWS fish data from 2001 through 2007 (using the "combined fish metric" described in Section 6.2.1.a) using multiple linear regression (MLR). As discussed above, this method was selected because it identified the habitat variables that statistically best explain the fish data, assigns relative weights to those variables to inform their relative importance, and produces a quantitative metric (the r-squared value) that can then be compared to the relative importance of other variables, such as water quality.

6.3.1 Interpretation of Best Subsets Multiple Linear Regression Results

The best subsets MLR method calculates regressions of all permutations of the independent variables (habitat) with the dependent variable (fish) and produces multiple regression models for inspection. The method does this by calculating a specified number of regression models using various numbers of variables from one up to the total number of variables. The MiniTab statistical software package was used to conduct the MLR analysis and it allows specification of the number of regression models produced in each variable set. For this study, the top three regression models were produced for each variable set. In other words, starting with a total of 15 variables, the analysis produced the top three regression models with one habitat variable, the top three regression models with two variables, and so on, up to 15 variables.

With multiple regression models calculated for each analysis, some means of discriminating between the regression models and for selecting a preferred model is needed. There are several factors that were considered in this study, when inspecting the MLR results:

- Number of variables – Because the best subsets MLR produced regression models with as few as one variable, and as many as 15, there was wide latitude in selecting regression models with a range of variable numbers. Although in some analyses the model with the fewest variables, all other things being equal, might be preferred, that was not the case here. The review of the regression models took into account the objectives of the study, specifically the need to support development of a descriptive index for physical habitat. In that sense, it can be argued that a greater number of variables is preferable to a fewer number of variables.

- Sign of the variables – Each variable that appears in a regression model has a positive or negative value. A positive value indicates that the habitat variable is positively correlated with the fish data and a negative sign indicates the opposite. In some cases, it was observed that variables intended to represent a positive habitat condition were assigned a negative sign in a particular regression model or vice versa. Due to the highly modified nature of the CAWS, this may have occurred in this study more than would occur in a study of natural systems. In any case, it may be counterproductive to use a regression that includes these variables. This is discussed in Section 6.3.2 below.

- R-squared and adjusted r-squared values – The r-squared value for each regression model was calculated and an "adjusted" r-squared was also calculated for each. The adjusted r-squared value accounts for the degrees of freedom in the regression. In other words, the raw r-squared value of the regression may be increased by adding more variables (degrees of freedom) but the statistical certainty of the calculated data relationship may be diminished. The adjusted r-squared value accounts for this and is, therefore, a truer measure of the regression model's descriptive ability. In comparing regression models, a higher adjusted r-squared was preferred.

- Mallow's C-p value – Mallow's C-p is a commonly used parameter in MLR analysis because it represents a measure of both the variance of the regression and the bias[7]. As more variables are added to the regression, C-p typically increases. Although a common interpretation of MLR results is to select the regression model with the lowest C-p (meaning the regression with the lowest total discrepancy (variance plus bias), such a model might not be the best fit to the data. A higher C-p value means a regression model with more discrepancies but, possibly, a better fit to the data. In general, a value of C-p that is equal to, or less than, the number of variables in the regression has the minimum bias. In comparing regression models in this study, a Mallow's C-p value less than the number of variables in the regression was preferred.

[7] In regression analysis, bias refers to the systematic overestimation or underestimation of the dependent variable by the regression model. This is different from variance, which is the natural variability or "scatter" of the variable.

- Variable confidence – For each variable included in each regression model, a statistical confidence level (p-value) was calculated. This value reflects the level of uncertainty in each variable and a 90% confidence level was preferred ($p<0.10$). Trade-offs between statistical certainty and regression fit were observed. Adding more variables might, in some cases, have increased the adjusted r-squared of the regression, but it might have diminished the statistical certainty of certain variables. The variable p-values were the last item to be examined and although the inclusion of variables with p-values greater than 0.1 did not automatically eliminate the regression from consideration, this factor was weighed.

All of these factors were considered when reviewing the MLR results in this study. In addition, the application of professional judgment and consideration of the objectives of the study were integral to the process. As stated in Draper and Smith (1981) when discussing selection of regression models, "all selection procedures are essentially methods for the orderly displaying and reviewing of data. Applied with common sense, they can produce useful results; applied thoughtlessly, and/or mechanistically, they may be useless or even misleading."

6.3.2 Discovery of Counterintuitive Variable Results

The initial MLR was conducted using available paired (concurrent and collocated) measurements of fish and habitat. In all, 81 paired fish/habitat "events" were used in this analysis. Initial MLR analyses presented some counterintuitive results for certain variables, described below:

- Flashiness appeared as a positively correlated variable with fish, when it generally is believed to be a negative condition reflecting watershed urbanization and increased imperviousness. It was concluded, given the highly regulated hydrology of the CAWS, that flashiness is not a truly meaningful habitat variable in the CAWS and that it's positive relationship to fish is an artifact of the data.

- The percent large substrate (gravel, cobbles, and boulders) in deep water appeared as both a negatively and positively correlated variable with fish, depending on which other habitat variable were used in a particular regression. This suggested a degree of instability and unreliability in the data for this variable.

- Similar to the percent large substrate in deep water, the variable representing the percentage of plant debris on the channel bottom appeared as both a positive and a negative variable in the different regressions. Again, this suggested a degree of instability and unreliability in the data for this variable.

Based on these observations, these three variables were eliminated from the regression analysis, so the final regressions between habitat variables and fish data were conducted using 12 habitat variables (Table 6-3).

Table 6-3: Final Habitat Variables Used in Multiple Linear Regression with Fish Data

Variable Category	Habitat Variable
Geomorphology & Hydrology	Wetted perimeter of channel Maximum depth in reach Number of off-channel bays Bank "pocket" areas
Sediment & Substrate	% Gravel, cobbles, boulders, shallow % Organic sludge
In-Stream Cover	Average macrophyte cover In-stream cover (present or absent)
Bank & Riparian Condition	Dominant riparian land use % vertical walled banks in reach % Riprap banks in reach
Anthropogenic Impacts	Manmade structures

6.3.3 System-Wide MLR Results

The MLR between the habitat variables and the combined fish metric was first run using the 2008 Secchi data, retroactively applied at each station for the 2001 – 2007 events. Using the best subsets method, the top three regression models for each possible number of variables were identified. Table 6-4 shows the results of this analysis. The habitat variables are listed across the top of the table and each row represents a different regression equation. The variables included in each regression are indicated by an "X" in the column for that variable.

The second and third columns present the r-squared and adjusted r-squared values for each regression. The r-squared is the basic "goodness of fit" measure, which indicates how much of the data variability is explained by the regression. An r-squared of 0.4 indicates that 40% of the data variability is explained by the regression equation. In general, the r-squared value will continue to increase as more variables are added, but there is a point beyond which the statistical reliability of the regression begins to diminish. To account for this, the adjusted r-squared is calculated, which takes into account the statistical reliability as a function of the number of variables, which is why the adjusted r-squared begins to decrease after a certain number of variables is reached.

Table 6-4: Summary of Regression Models for System-Wide Comparison of Fish and Habitat Data for 2001 – 2007

No. Vars	R-squared	Adjusted r-squared	Mallows C-p	WET_PER	MAX_DEP	OFF_CH_BAY	BANK_POC_AREA	BIG_S	CAWS_ORGSLG	DOM_LU	BNK_WALL	BNK_RIPRAP	MAN_MADE_STRUC	MCRPH_CHAN	NUMCOV
1	0.25	0.24	25.2		X										
1	0.15	0.14	38.6						X						
1	0.15	0.14	39.0											X	
2	0.35	0.34	12.8		X							X			
2	0.33	0.31	16.2		X	X									
2	0.31	0.29	19.3		X								X		
3	0.42	0.4	6.1		X				X			X			
3	0.4	0.37	9.3		X							X	X		
3	0.4	0.37	9.4		X						X	X			
4	0.44	0.41	5.4		X							X	X		
4	0.44	0.41	5.5		X						X	X	X		
4	0.43	0.40	6.1		X				X		X	X		X	
5	0.47	0.43	3.5	X	X							X	X	X	
5	0.47	0.43	3.6		X						X	X	X		
5	0.46	0.42	4.9		X	X					X	X	X		
6	0.48	0.44	3.6	X	X					X		X	X	X	
6	0.48	0.44	4.0		X	X					X	X	X	X	
6	0.48	0.43	4.2	X	X						X	X	X	X	
7	0.49	0.44	4.0	X	X					X		X	X	X	X
7	0.49	0.44	4.6	X	X					X	X	X	X	X	
7	0.49	0.44	5.0	X	X			X		X		X	X	X	
8	0.5	0.44	5.2	X	X			X		X		X	X	X	X
8	0.5	0.44	5.5	X	X			X		X	X	X	X	X	
8	0.49	0.44	5.9	X	X		X			X		X	X	X	X
9	0.5	0.44	7.0	X	X			X		X	X	X	X	X	X
9	0.5	0.44	7.2	X	X			X	X	X		X	X	X	X
9	0.5	0.44	7.2	X	X		X	X		X		X	X	X	X
10	0.5	0.43	9.0	X	X	X		X		X	X	X	X	X	X
10	0.5	0.43	9.0	X	X			X	X	X	X	X	X	X	X
10	0.5	0.43	9.0	X	X		X	X		X	X	X	X	X	X
11	0.5	0.42	11.0	X	X	X	X	X		X	X	X	X	X	X

As shown in Table 6-4, the regression models have adjusted r-squared values ranging from 0.14 to 0.44. The regression models with four variables or fewer have lower adjusted r-squared values and C-p values that are greater than the number of variables, indicating relatively high bias (systematic overestimation or underestimation of the data), so these were not considered further. The maximum adjusted r-squared value of 0.44 was achieved with regression models having six or more variables. Increasing the number of variables beyond six did not increase the adjusted r-squared value, but increased the C-p values and also resulted in some significantly increased P-values (not presented in the table), suggesting there was little benefit to using a regression model with more than six variables.

The two 6-variable regression models having adjusted r-squared values of 0.44 contained five variables in common. One regression model included channel wetted perimeter as the sixth variable and the other included off-channel bays as the sixth variable. With this as the point of comparison, the model including off-channel bays was selected because this variable was more intuitively understandable in terms of its habitat benefit than channel wetted perimeter.

The six-variable regression that is selected from this process included the following habitat variables:

- Maximum depth of channel ($p=0.000$)
- Off-channel bays ($p=0.197$)
- Percent of vertical wall banks in reach ($p = 0.053$)
- Percent of riprap banks in reach ($p = 0.001$)
- Manmade structures in reach ($p = 0.019$)
- Percent macrophyte cover in reach ($p = 0.086$)

The regression calculated using these variables had a raw r^2 of 0.48 and an adjusted r^2 of 0.44. This result indicated that the six variables in the regression account for 48% of the variability in the fish data in the CAWS. The equation for this regression was:

$$\begin{aligned} CFM = 12.8 - 0.381 \times MAX\_DEP + 1.03 \times \ln(OFF\_CH\_BAY + 1) - 2.03 \times \\ \operatorname{asin}((BNK\_WALL)^{0.5}) - 1.11 \times (\ln(BNK\_RIPRAP + 1)) - 6.06 \times \\ \ln(MAN\_MADE\_STRUC + 1) + 0.214 * MCRPH\_CHAN \end{aligned}$$

Where:

CFM = Combined fish metric

MAX_DEP = The maximum channel depth in reach

OFF_CH_BAY = the number of areas in the reach that function as off-channel bays, providing refuge for fish

BNK_WALL = the percentage of bank, by length, occupied by vertical walls

BNK_RIPRAP = the percentage of riprap banks in reach, by length

MAN_MADE_STRUC = the number of manmade structures in the reach

MCRPH_CHAN = the percentage macrophyte cover in the reach.

Each of the variables in this regression has a p-value less than 0.1, which represents 90% confidence, except off channel bays, which has a p-value of 0.197 (~80% confidence). A plot depicting this regression is presented in Figure 6-3.

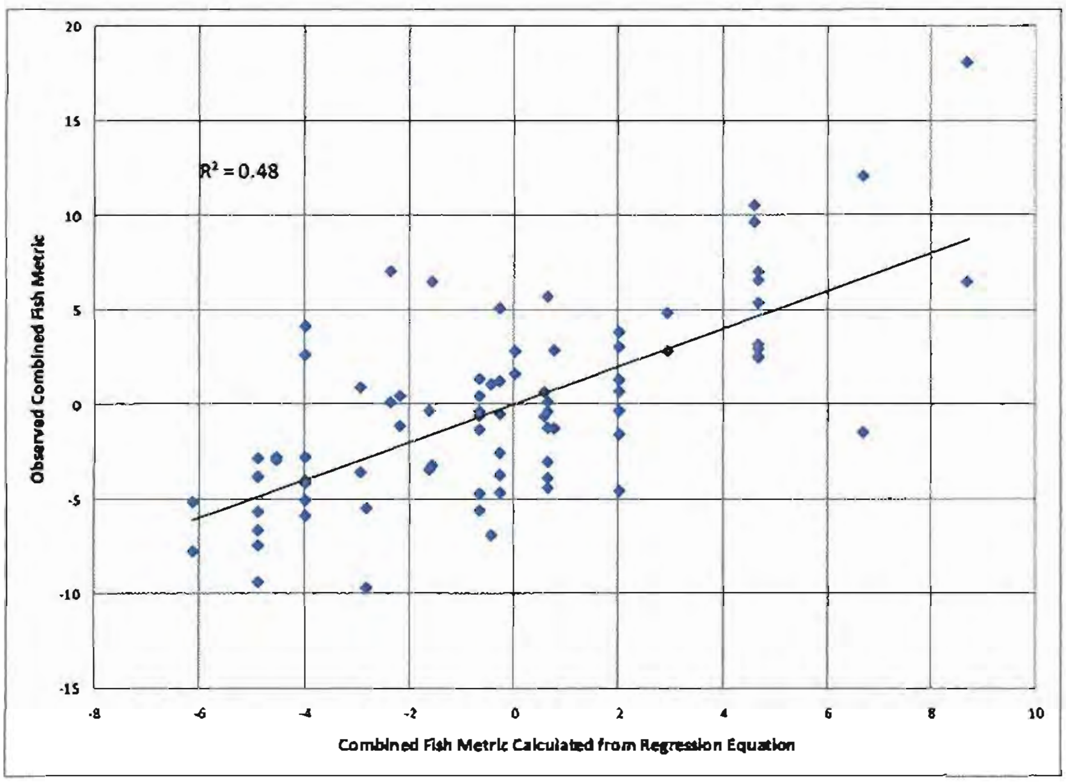

Figure 6-3: Plot of CAWS Six-Variable Habitat Regression Model with 2001-2007 Fish Data.

One of the underlying assumptions of MLR is that the regression residuals (predicted values minus observed values) follow the normal distribution. The normal probability plot depicted in Figure 6-4 shows that the residuals are normally distributed. Values in a normal distribution will fall on the diagonal line.

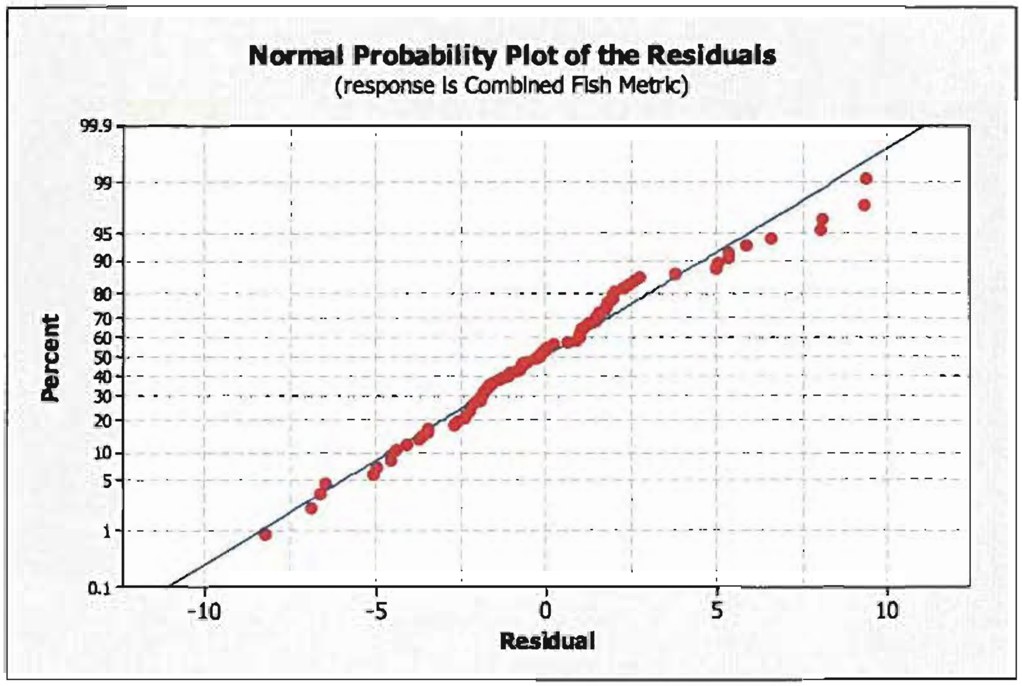

Figure 6-4: Normal Probability Plot of Regression Residuals for the Selected Six-Variable CAWS Habitat Regression with Fish Data.

In addition to the assumption of normality, it is assumed that the residuals are independent. This is evaluated using a scatter plot of residuals against fitted values, as depicted in Figure 6-5.

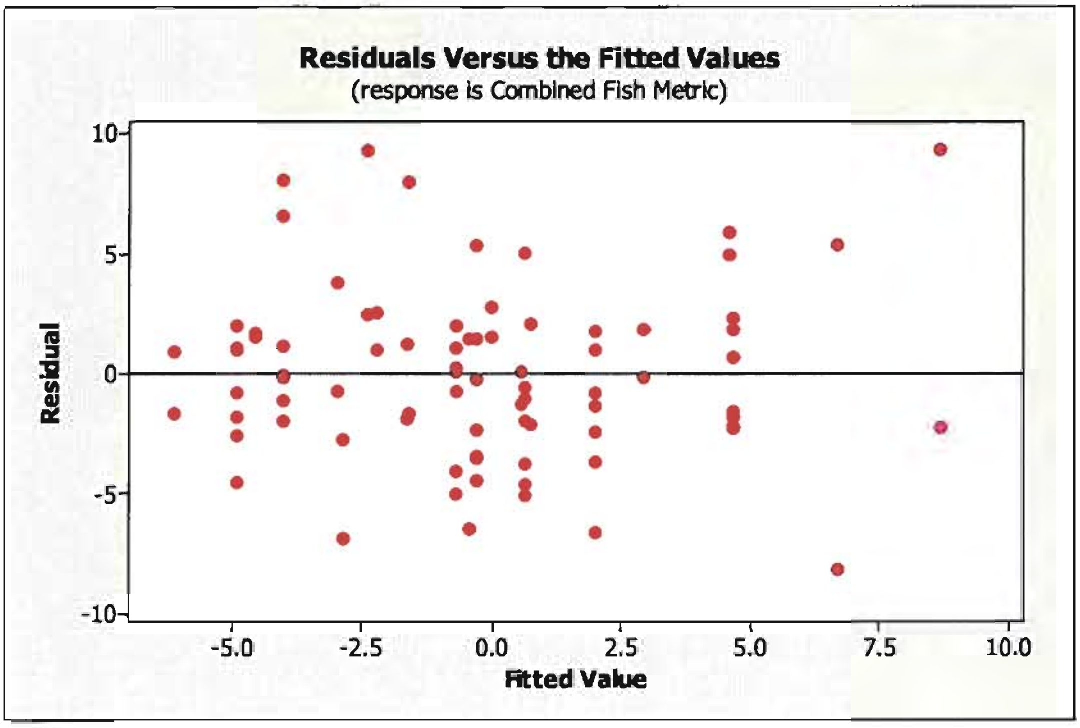

Figure 6-5: Scatter Plot of Regression Residuals vs. Fitted Values for the Six-Variable CAWS Habitat Regression.

The values of the residuals plotted against the fitted value appear to be randomly distributed, suggesting that the residuals are independent. Based on these analyses of the regression residuals, the seven-variable CAWS habitat regression appears to uphold the underlying assumptions of normality and independence.

6.3.4 Comparison of Habitat Regressions to 2008 Fish Data

To evaluate and verify the usefulness of the regression model described above, 2008 fish data were used. In 2008, fish samples were collected at 20 stations in the CAWS Study area, which included 14 stations sampled by the District and six supplemental stations sampled by LimnoTech and their subcontractor Ecological Specialists, Inc. The combined fish metric for these 20 stations was calculated from the 2008 fish data and compared to the habitat regression model described above, calculated at the 20 stations. Comparison of the six-variable regression model to the 2008 fish data is depicted graphically in Figures 6-6.

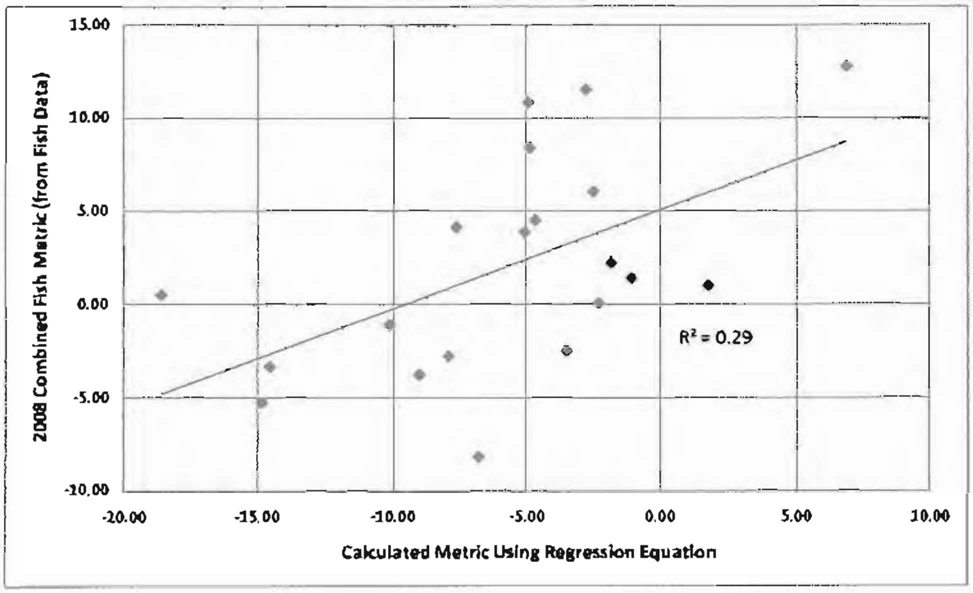

Figure 6-6: Comparison of the CAWS Habitat Regression Model with 2008 Fish Data.

As shown in Figure 6-6, the six-variable habitat regression model (developed using 2001 – 2007 fish data) shows a relatively good fit with the 2008 fish data. The r-squared value of 0.29 ($p = 0.014$) indicates that there is good and statistically significant correlation (98.6% confidence) between the habitat regression model and the 2008 fish data.

It is also of interest to know how this regression might correlate with long-term averages in CAWS fisheries condition. To evaluate this, the average combined fish metric at each CAWS sampling station was calculated from the 2001 – 2008 data and the regression equation was compared to these averages. Figure 6-7 shows this comparison. The regression fit the long-term averages with an r-squared of 0.51, indicating that the six habitat variables in the regression equation explain more than 50% of the variability in fish data over long periods.

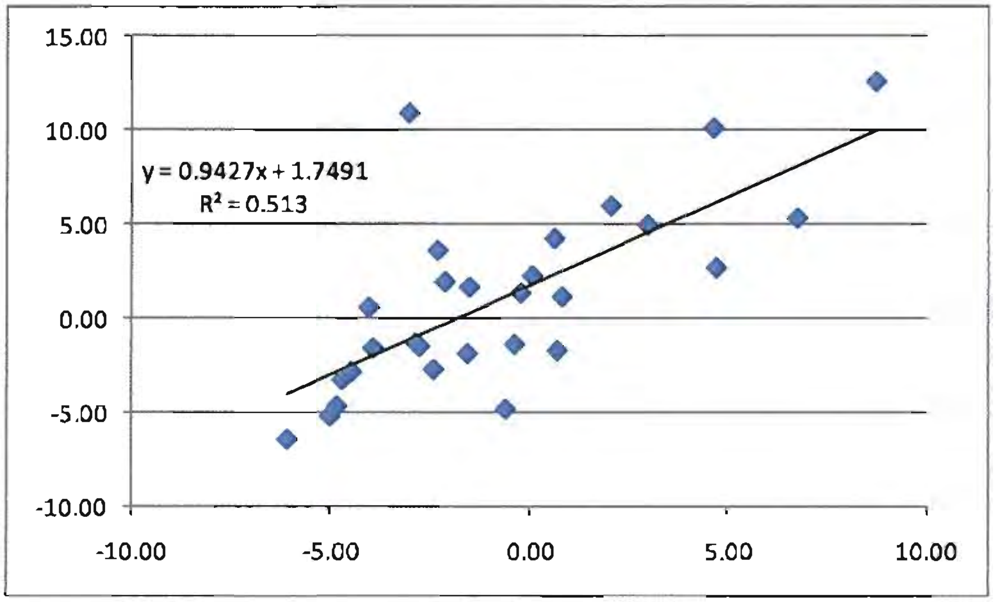

Figure 6-7: Comparison of the CAWS Habitat Regression Model with Averaged Fish Data (2001 – 2008).

This comparison is further verification of the importance of the six habitat variables in the habitat regression and indicates that the regression can provide a solid foundation for development of a habitat index for the CAWS.

6.4 RELATIVE IMPORTANCE OF PHYSICAL HABITAT IN THE CAWS

The regression analysis of physical habitat with fish can be used to evaluate the relative importance of habitat to fish in the CAWS. As previously discussed, the regression analysis shows that physical habitat can explain 48% of the fish data collected from 2001 – 2007. While this is a significant finding, it means that approximately half of the fish data is not explained by the six habitat variables in the regression. The following sections evaluate what else might be contributing to variability in CAWS fish data.

6.4.1 Variation in Fish Data Not Explained by Habitat Variation

The observation that physical habitat conditions can explain up to approximately half of the variability in fish data raises the question as to what can explain the rest of the variability in CAWS fish data. To investigate this, two evaluations were performed using the regression residuals:

- The regression residuals were compared to the station-by-station variation in fish data between the 2001-2007 dataset and the 2008 dataset. This comparison was performed to evaluate how much of the unexplained variability in fish data may be attributable to variation in fish over time.

- The regression residuals were compared to DO metrics at each station. This comparison was performed to evaluate how much of the variability in fish data, not explained by the key habitat variables represented in the regression equation, may be attributable to DO.

The regression equation used for these comparisons was the six-variable regression equation presented in Section 6.3.3. These comparisons are depicted graphically in Figures 6-8 and 6-9, respectively.

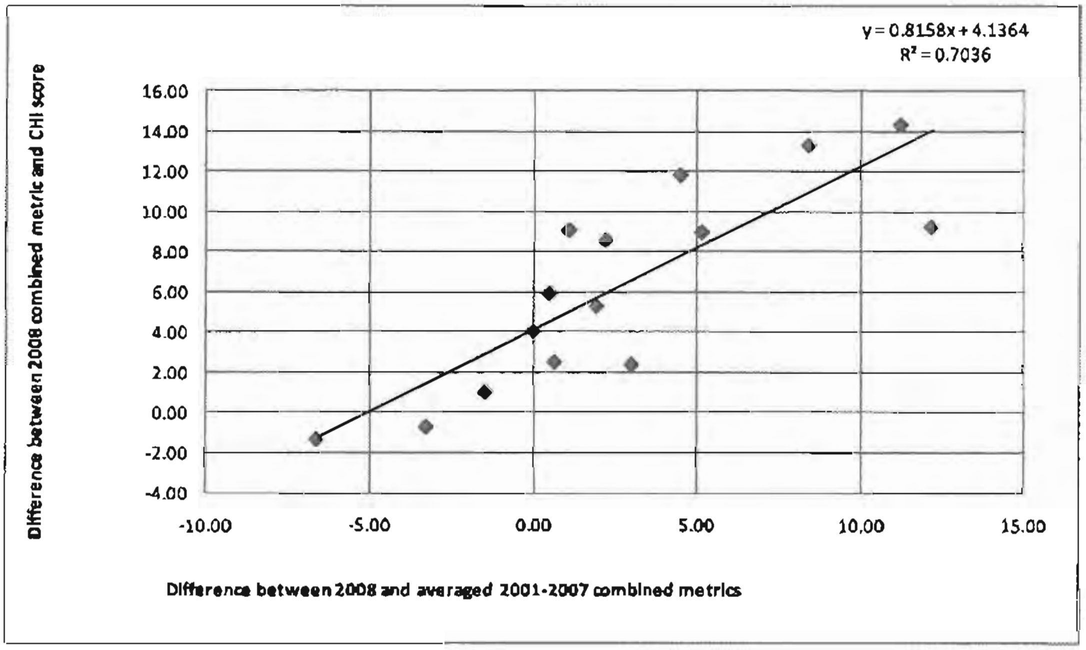

Figure 6-8: Comparison of Regression Residuals with Variation in Metrics Calculated Using Fish Data from 2001-2007 and 2008.

Figure 6-8 compares the habitat regression residuals (predicted values minus observed values) to the difference between the average fish metric values for the 2001 to 2007 data period (used for regression development) and the 2008 data set (used for regression validation). This comparison shows a relatively strong correlation (r-squared = 0.70) between the regression residuals and the change in fish metrics from the 2001-2007 period and 2008. This suggests that as much as 70% of the variability in the CAWS fish data that is not explained by the six habitat variables in the regression equation (35% of total variability in fish data) can be explained by variability in the fish samples themselves, as opposed to some other external condition, such as a missing habitat variable.

To further investigate this, the error associated with year-to-year variability of the combined fish metric at individual sampling stations was compared to the error of the regression model. Table 6-5 shows the standard deviation of the CFM at each of the stations. The mean standard deviation (the square root of variance from the mean) of the CFM measurements is 3.1 while the regression model root mean squared error (the square root of variance from the predicted value) is 3.7. The fact that the mean standard deviation is 3.1, which is nearly equal to the root mean squared error of 3.7, suggests that hat suggests that the majority of the model error is due to the year-to-year variability of the fish measurements.

Table 6-5: Standard Deviation of the Combined Fish Metric at District Sampling Stations.

Station_No	Station Name	n	Mean CFM	St Dev
1014	North Shore Channel at Central Street	2	12.3	8.2
1015	North Shore Channel at Touhy Avenue	7	0.3	2.9
1016	North Branch Chicago River at Wilson Avenue	2	-1.4	3.2
1017	South Branch Chicago River at Madison Street	2	3.6	4.9
1018	Chicago Sanitary and Ship Canal at Damen Avenue	2	-0.4	1.1
1019	Chicago Sanitary and Ship Canal at Harlem Avenue	7	-1.3	3.5
1020	Chicago Sanitary and Ship Canal at Route 83	2	1.6	6.9
1021	Calumet-Sag Channel at Route 83	2	-6.5	1.8
1022	North Branch Chicago River at Grand Avenue	7	-1.0	3.4
1023	Chicago Sanitary and Ship Canal at Stephen Street	2	-7.6	3.0
1029	Little Calumet River at Indiana Avenue	2	0.8	2.9
1031	Calumet-Sag Channel at Ashland Avenue	2	-2.9	0.1
1032	Calumet-Sag Channel at Cicero Avenue	7	-1.6	2.6
1034	North Branch Chicago River at Diversey Parkway	2	-2.9	5.6
1035	Chicago River at Lake Shore Drive	2	10.1	0.6
1036	Chicago Sanitary and Ship Canal at Cicero Avenue	7	-2.2	3.9
1037	Little Calumet River at Halsted Street	7	4.3	2.0
1045	Chicago Sanitary and Ship Canal at Lockport (16th Street)	7	-5.7	2.3
1048	Bubbly Creek at Archer Avenue	2	0.0	0.9
1049	Chicago River at Wells Street	2	2.2	0.8
1050	North Shore Channel at Foster Avenue	2	3.8	1.4
1051	North Shore Channel at Oakton Street	2	5.3	9.6
1056	South Branch Chicago River at Loomis Street	2	-1.9	2.2
	Mean across all stations (weighted by number of samples)	**81**	**-0.2**	**3.1**

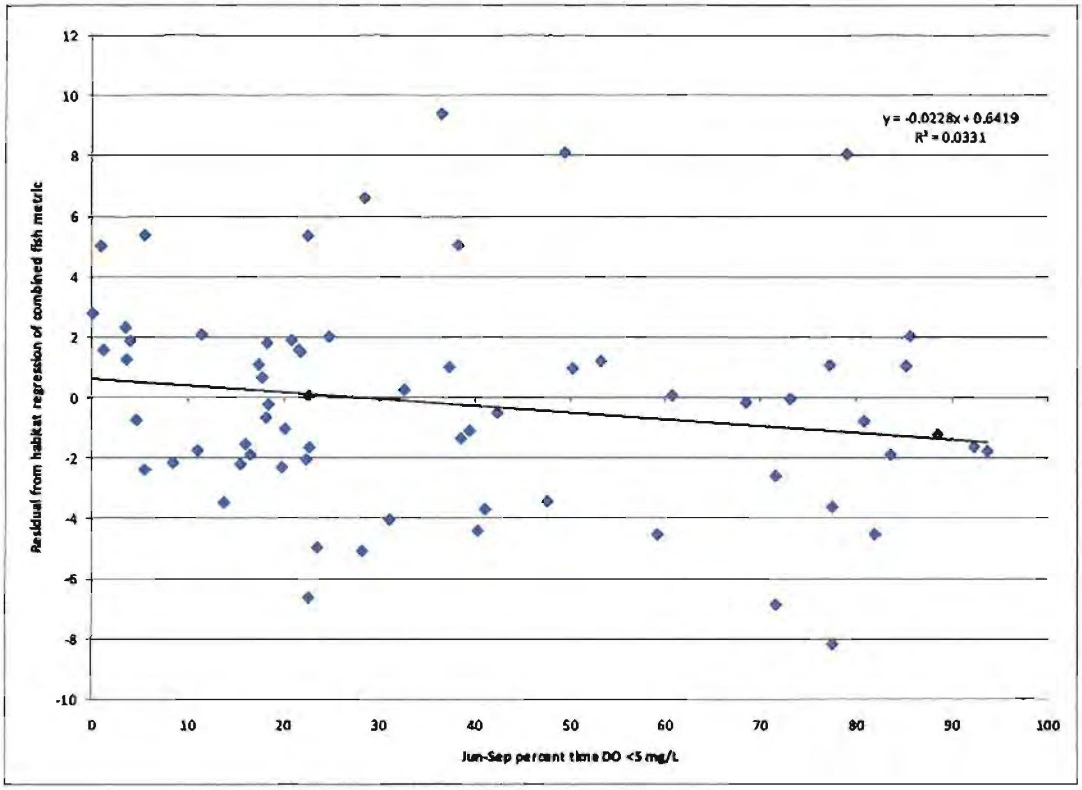

Figure 6-9: Comparison of Regression Residuals with Percent of Time Dissolved Oxygen Less Than 5 mg/L.

Figure 6-9 compares the habitat regression residual to the percent of time that DO was less than 5 mg/L at each station from June through September. This water quality metric was found to be the most highly correlated with individual fish metrics in the CAWS, as reported in Appendix C. The regression has an r-squared = 0.03, which indicates that only 3% of the CAWS fish data variability that is not explained by the six habitat variables in the regression equation (1.5% of total variability in fish data) may be explained by DO conditions at each sampling station.

6.4.2 Relative Importance of Habitat Versus Water Quality in the CAWS

The regression analysis presented in Section 6.3.3 shows that physical habitat alone can explain up to 48% of fish data collected in the CAWS from 2001 – 2007, which is significantly better than can be accomplished by evaluating water quality alone. In the analysis presented in Appendix C, the DO metric most highly correlated with fish data only had an r-squared of 0.27, meaning that DO alone can only explain 27% of the variability in the same seven years of fish data. This indicates that physical habitat is relatively more important in understanding fisheries in the CAWS than water quality.

To further investigate the relative importance of physical habitat and water quality to fish in the CAWS, A key DO metric (the percent of time that DO is less than 5 mg/L at each station from June through September) was added to a key habitat regression discussed above, to observe whether the inclusion of the DO variable would significantly improve the ability of the regression equation to explain the fish data.

It should be noted that a wide range of water quality metrics were evaluated with respect to fish data, to identify the metric most correlated to fish metrics, which was the percent of time that DO is less than 5 mg/L at each station from June through September. The six-variable regression equation discussed in section 6.3.3 was used for this test. That regression equation, developed using system-wide data, included the following variables:

- Maximum depth of channel
- Off-channel bays
- Percent of vertical wall banks in reach
- Percent of riprap banks in reach
- Manmade structures in reach
- Percent macrophyte cover in reach

The percent of time between June and September that DO was below 5 mg/l was added to this set of habitat variables because it was the water quality variable identified as having the strongest relationship to fish in the CAWS. This set of variables was then compared to fish data using multiple linear regression. It should be noted that this regression was conducted on a slightly smaller dataset (67 events) because continuous DO data were not available at all of the CAWS stations with fish and habitat data.

In the original regression using habitat variables alone, the comparison to fish data yielded an r-squared value of 0.48, meaning that the habitat variables explained about 48% of the fish data. With the reduced data set, the r-squared dropped to 0.42, probably because fewer data were used. When DO was added to the variable set and a new regression was calculated, the r-squared of the new regression with fish data was 0.46. This result indicates that including DO with the habitat variables improved the amount of fish data variability explained by the regression by about 4% over physical habitat alone.

6.4.3 Summary Findings for Relative Importance of Habitat in the CAWS

From these comparisons and the overall analysis of the relationship of physical habitat to fish in the CAWS, the following conclusions can be made:

- The two most important physical habitat variables in the CAWS that are positively correlated with fish are the amount of macrophyte cover and the quantity of areas that act as off-channel bays to provide refuge from the main channel.

- The four most important physical habitat variables in the CAWS that are negatively correlated with fish are the maximum depth of the channel, the amount of vertical walled banks, the amount of riprap banks, and the number of manmade structures.

- These six variables account for 48% (approximately half) of the variability in fish data collected in the CAWS from 2001 – 2007.

- Of the half of fish data variability that is not explained by these physical habitat variables, as much as 70% of that half can be explained by variation in fish sampling results from year to year. This means that the fish measured at a location can vary significantly from one sample event to the next and that this will lead to an inherent variability in the data that cannot be explained by changes in independent variables such as habitat or water quality.

- The percent of time that DO is less than 5 mg/L at a given station in the CAWS from June through September explains approximately 3% of the half of the fish data variability that is not explained by the six key physical habitat variables.

- DO is much less important to fish in the CAWS than physical habitat. DO alone can only explain between 2% and 27% of the fish data variability, while the physical habitat can explain 48%. The addition of the key DO metric to the main habitat variables only resulted in a 4% improvement over using habitat alone.

The use of these findings in developing a CAWS-specific habitat index is discussed in the next section.

This page is blank to facilitate double sided printing

7. DEVELOPMENT OF A CAWS HABITAT INDEX

The process outlined in Section 6 of this report systematically narrowed the field of potentially important habitat variables from 241 original variables to a final set of six habitat variables that represent the most statistically important measured habitat variables to fish in the CAWS. These six variables are:

- Maximum depth of channel
- Off-channel bays
- Percent of vertical wall banks in reach
- Percent of riprap banks in reach
- Manmade structures in reach
- Percent macrophyte cover in reach

Together, these habitat variables explain 48% of the fish data variability in the CAWS. The development of a system-specific habitat index is discussed in this section, with emphasis on the following topics:

- Objectives for the CAWS Habitat Index (Section 7.1) – The main objectives for a system-specific CAWS habitat index are outlined in this section.
- Use of the CAWS Habitat Regression Equation (Section7.2) – This section discusses the role of the CAWS habitat regression in developing a habitat index for the system.
- CAWS Habitat Index Development (Section 7.3) – Development of a CAWS-specific habitat index is discussed.
- Potential Limitations of the CAWS Habitat Index (Section7.4) – Potential limitations of the CAWS habitat index presented in Section 7.3 are described.

7.1 OBJECTIVES FOR THE CAWS HABITAT INDEX

One of the original objectives for this study, as discussed in Section 1 was to "use a multi-metric habitat index to evaluate physical habitat conditions in the CAWS and use physical habitat data and the above multi-metric index to assess the relative importance of physical habitat to fish in the CAWS." As discussed in Section 2, no existing habitat indices for non-wadeable waters were identified that would be applicable to the CAWS, therefore development of a system-specific index would be required. The process of developing a system-specific habitat index required detailed, in-depth analysis of habitat and fish data. This process of data analysis, while paving the way for development of a system-specific habitat index for the CAWS, was also sufficient to meet the objectives for which the index was originally thought to be needed. Specifically, the evaluation of physical habitat conditions in the CAWS and the assessment of the relative importance of physical habitat to fish in the CAWS was addressed without an index, as discussed in Section 6.

As such, the objectives for a habitat index for the CAWS have shifted somewhat from what was originally envisioned. With the completion of the analysis documented in this report, the objectives for a CAWS-specific habitat index should be to:

- Provide a tool for characterization of reaches within the CAWS for purposes of comparing the range of habitat quality within the CAWS and for prioritizing locations for potential habitat improvement measures.
- Provide a tool for characterizing habitat changes in reaches over time.
- Represent the habitat attributes that are most important to aquatic biota in the CAWS, based on system-specific data.

The technical literature on the subject present different approaches for developing habitat indices and a single, universally accepted standard method has not been identified. The flowing sections address the use of the multiple linear regression analyses discussed previously in developing a CAWS-specific habitat index.

7.2 USE OF THE CAWS HABITAT REGRESSION EQUATION

One method for using the habitat regression presented in Section 6 to develop a CAWS-specific habitat index is to use a regression equation directly as an index equation to measure habitat quality in the CAWS. This has certain advantages, including the fact that the index would only include the habitat variables that are currently most important to the biotic indicator population (fish in this Study). Direct use of the variable coefficients from the regression equation as weights for the variables in the index would be the most statistically sound approach.

However, this approach has a significant limitation, in that it can ignore other important habitat variables that can be used to characterize physical habitat in the system. Using only variables from the regression analysis may omit variables that are important, but not as relatively important as those in the regression. For example, overhanging riparian vegetation was not included in the final habitat regression because it was highly correlated with vertical walled banks. This does not mean that it is not an important habitat variable. The bank pocket area variable was included in the regression analysis, but did not appear in the selected regression. This does not mean that these small bank refuges are unimportant to fish. A better approach is to use the regression analysis to inform the habitat index by pointing to important variables and by helping understand the relative importance of those variables. This allows for the application of professional judgment, informed by knowledge of the system, the data, and aquatic ecology in general. This approach is described below.

7.3 CAWS HABITAT INDEX DEVELOPMENT

As stated at the beginning of this section, the regression analyses presented in Section 6 identified six physical habitat variables that are the most important to fish in the CAWS, based on the data and analytical methods used in this study. Because they are the most important variables for understanding habitat quality, they are the best candidates for a CAWS-specific habitat index. In addition, other habitat variables were not included in the selected regression, but were evaluated for inclusion in the CAWS habitat index, as discussed below.

To evaluate the effect of including additional variables with the selected regression equation as the basis for an index, an index development spreadsheet was created using the regression equation, which would allow comparison of the regression calculation to the average combined fish metric at each station, for the monitoring period used in this study (2001-2008). This comparison was depicted graphically in Figure 6-6 and shows that the regression equation versus the average combined fish metric for each station has an r-squared of 0.51, meaning that the regression can explain 51% of the variability in long-term average fish data in the CAWS.

The index development spreadsheet also included station-by-station values of the following other habitat variables of interest:

- Bank pocket areas – This variable was used in the regression analysis but does not appear in the selected regression. It represents the count of relatively small bank refuge areas for fish and was included because it can represent an important cover variable.

- Large substrate in shallow and deep parts of the channel – These variables were also included in the regression analysis but did not appear in the selected regression. They were considered in the index development because of the general importance of large substrate to fish.

- Organic sludge – This variable was included in the regression analysis but did not appear in the selected regression. It represents a general substrate condition in some of the CAWS reaches that indicates very fine sediment with residual impacts of industrial chemicals. It was included because it may be an important local limitation to ecological health in parts of the CAWS.

- Overhanging vegetation – Overhanging riparian vegetation is recognized as important in aquatic systems for providing shade and a source of organic material and food (insects) for some fish. This variable was not included in the habitat regression analysis because it is strongly correlated with another variable, vertical wall banks. It is, however, an important habitat variable and should be included in the index.

The regression equation is simply the sum of the values for each included variable, each multiplied by a coefficient. The coefficients in the regression equation are determined by the statistical process. In adding variables to this equation, the assumption was made that none of the additional variables is more important than the variables in the original regression equation; otherwise they themselves would have appeared in the equation. Therefore, it was assumed that none of the additional variables could have a larger coefficient than the lowest coefficient already in the regression equation. In other words, the additional variable could not be weighted more heavily than a variable that appeared in the regression.

It was also recognized that the addition of variables would degrade the fit of the equation to the data. For index development, the average combined fish metric at each station was calculated for the 2001 – 2008 period. As described above and in Section 6.3.4, the regression equation had an r-squared of 0.51 with these long-term averages. It would be expected that adding variables to the equation would result in a lower r-squared, so there is a trade-off between adding variables and the r-squared value. It was decided that the addition of variables to the regression equation should not result in an r-squared less than 0.48, which was the r-squared that the original regression had with the 2001-2007 data, when it was originally developed.

With these constraints, the additional variables were tested alone and in combination, using coefficients less than 0.2, which was the lowest coefficient assigned to a variable in the original regression. Using this approach, a combination of coefficients was developed that matched the r-squared of the original regression (0.48). The variables and their coefficients are listed in Table 7-1. The variable values used in this analysis are presented in Table 7-2.

Table 7-1: Habitat Variables and Coefficients Used in CAWS Habitat Index.

Habitat Variable	Coefficient
Maximum depth of channel (-)	0.381
Off-channel bays (+)	1.03
Vertical wall banks (-)	2.03
Riprap banks (-)	1.11
Manmade structures (-)	6.06
Macrophyte cover (+)	0.214
Overhanging vegetation (+)	0.1
Bank pocket areas (+)	0.05
Large substrate - shallow (+)	0.005
Large substrate - deep (+)	0.005
Organic sludge (-)	0.08

Table 7-2: Values of Habitat Variables Assigned to CAWS Stations for Index Development.

Reach	Maximum Channel Depth (ft)	Off-Channel Bays	Vertical Wall Banks (%)	Riprap Banks (%)	Manmade Structures	Macrophyte Cover (%)	Overhanging Vegetation (%)	Bank Pocket Areas	Large Substrate – Shallow (%)	Large Substrate – Deep (%)	Organic Sludge (%)
AWQM 35 - Upper North Shore Channel	8	2	0	0	1.0	9	33	0	20	0	0
AWQM 102 - Lower North Shore Channel	6	1	0	0	2.0	10	29	3	0	0	0
AWQM 36 - Lower North Shore Channel	12	3	7	22	1.8	13	33	3	42	8	0
AWQM 101 - Lower North Shore Channel	10	3	5	6	2.0	9	29	8	25	0	0
AWQM 37–No. Branch Chicago River No. of Addison	12	1	0	100	2.0	0	25	15	85	0	0
AWQM 7 - No. Branch Chicago River So. of Addison	12	3	19	81	1.0	0	10	9	5	0	0
AWQM 46 -No. Branch Chicago River So. of Addison	13	7	100	0	1.7	0	0	2	9	3	9
AWQM 74 - Chicago River (Lake Shore Drive)	8	7	60	0	2.5	10	0	10	5	0	0
AWQM 100 - Chicago River (Wells St.)	21	8	97	0	1.0	0	0	0	19	0	6
AWQM 39 - South Branch Chicago River	23	9	100	0	1.5	0	0	6	0	0	8
AWQM 108 - South Branch Chicago River	22	4	77	0	1.5	0	0	4	3	18	4
AWQM 99 - Bubbly Creek	13	1	35	0	2.0	0	8	9	5	5	48
AWQM 40 - Chicago Sanitary and Ship Canal	20	3	67	0	2.0	0	0	6	53	0	38
AWQM 75 - Chicago Sanitary and Ship Canal	19	3	13	23	2.2	1	4	16	35	5	6
AWQM 41 - Chicago Sanitary and Ship Canal	20	5	48	0	1.8	3	3	10	80	13	5
S1 - Chicago Sanitary and Ship Canal	19	8	0	20	4.0	6	3	12	5	0	1
S2 - Chicago Sanitary and Ship Canal	24	1	100	0	1.0	0	14	4	0	0	0
AWQM 42 - Chicago Sanitary and Ship Canal	25	1	100	0	0.5	0	11	19	0	0	0
AWQM 48 - Chicago Sanitary and Ship Canal	26	4	100	0	1.0	0	2	20	3	0	0
AWQM 92 - Chicago Sanitary and Ship Canal	26	4	52	4	1.7	1	3	6	34	3	8
AWQM 43 - Cal-Sag Channel	18	0	51	49	2.0	0	5	8	25	0	13
S3 - Cal-Sag Channel	14	2	0	50	2.0	0	5	17	20	0	0
S4 - Cal-Sag Channel	15	2	8	48	3.0	0	4	10	70	2	5
S5 - Cal-Sag Channel	14	3	19	49	3.0	0	7	13	0	10	10
AWQM 58 - Cal-Sag Channel	15	0	49	51	1.5	0	3	10	25	3	18
AWQM 59 - Cal-Sag Channel	15	1	49	0	2.2	0	5	14	7	18	11
AWQM 56 - Little Calumet River	16	3	2	14	1.0	1	5	20	1	21	9
AWQM 76 - Little Calumet River	14	8	0	0	1.3	1	8	14	10	31	0

These 11 variables represent a good mix of habitat variables including bank condition, in-stream cover, substrate, and anthropogenic impact. They also represent variables that are relatively easy to measure and many may be alterable to improve habitat in the future. The equation for the raw CAWS habitat index is:

$$CHI = 12.8 - 0.381 \times MAX\_DEP + 1.03 \times \ln(OFF\_CH\_BAY + 1) - 2.03 \times \operatorname{asin}((BNK\_WALL)^{0.5}) - 1.11 \times (\ln(BNK\_RIPRAP + 1)) - 6.06 \times \ln(MAN\_MADE\_STRUC + 1) + 0.214 * MCRPH\_CHAN + 0.1 \times PER\_COV\_ALT + 0.05 \times BANK\_POC\_AREA + 0.005 \times BIG\_S + 0.005 \times BIG\_D - 0.08 \times CAWS\_ORGSLG$$

Where:

CHI = raw CAWS Habitat Index

MAX_DEP = The maximum channel depth in reach

OFF_CH_BAY = the number of areas in the reach that function as off-channel bays, providing refuge for fish

BNK_WALL = the percentage of bank, by length, occupied by vertical walls

BNK_RIPRAP = the percentage of riprap banks in reach, by length

MAN_MADE_STRUC = the number of manmade structures in the reach

MCRPH_CHAN = the percentage macrophyte cover in the reach

PER_COV_ALT = the percent overhanging vegetation

BANK_POC_AREA = the number of bank pocket areas

BIG_S = the percentage of large substrate (gravel, cobbles, boulders) in the shallow part of the channel

BIG_D = the percentage of large substrate (gravel, cobbles, boulders) in the deep part of the channel

CAWS_ORGSLG = the percentage of organic sludge in sediment samples

The index values calculated for each CAWS sampling station from 2001 – 2008 are graphically compared to the average combined fish metric at those stations in Figure 7-1. It should be noted that in the index development stage, the raw values of the index calculation were used. The final index is normalized to a 0 to 100 scale, as explained in Section 7.4.

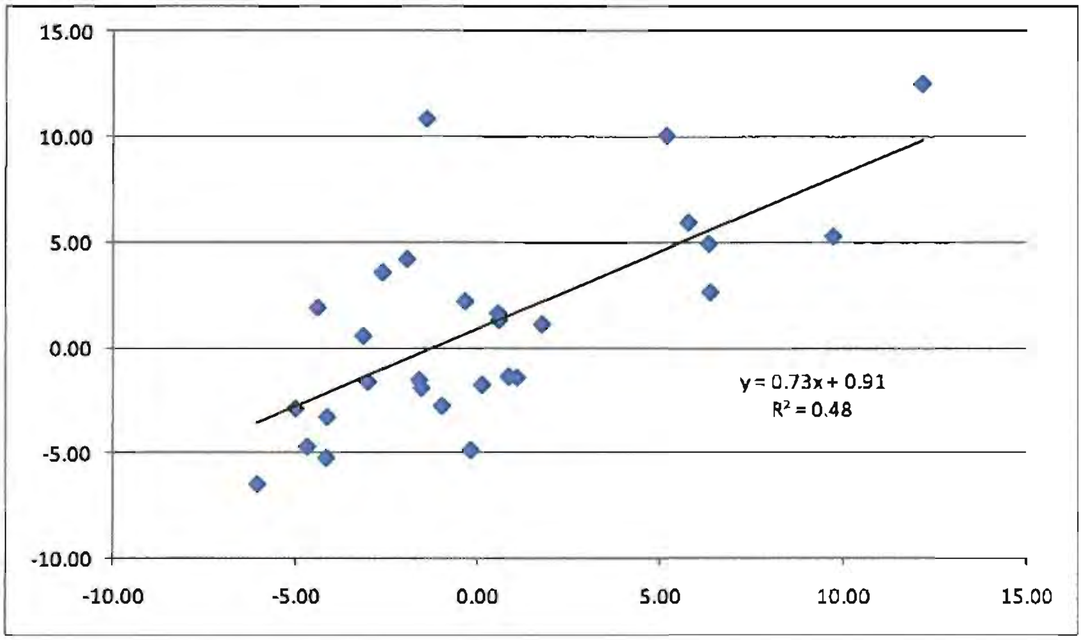

Figure 7-1: CAWS Habitat Index Compared to Average (2001-2008) Combined Fish Metric for Each Sampling Station.

As mentioned above, the r-squared of the CAWS habitat index to the fish data maintains the goodness of fit that the original habitat regression had, but it also compares well with comparisons reported for other habitat index studies that used multiple linear regression, as shown in Table 7-3.

Table 7-3: Comparison of Regression Coefficient Used in CAWS Habitat Index Development with Other Habitat Indices.

Habitat Index	Regression Coefficient for Index Development	Reference
CAWS	0.48	-
QHEI: comparison to IBI	0.45	Rankin, 1989
Maryland Physical Habitat Index: comparison to IBI	0.52	Hall et al., 1999
MI Non-Wadeable Habitat Index: comparison to catchment and riparian disturbance gradients	0.34/0.73	Wilhelm et al., 2005

The application of this index to individual reaches in the CAWS is presented in Section 7.4.

7.4 APPLICATION OF HABITAT INDEX BY REACH

The CAWS habitat index was calculated for each station as part of the index development, but it may also be useful for evaluating and comparing entire reaches in the CAWS. To do this, representative values had to be determined for each of the major reaches. The basis for assigning values of each variable is summarized in Table 7-4. The values assigned to each reach for each variable are presented in Table 7-5.

Table 7-4: Basis for Determining Reach-Wide Values of Key Habitat Variables.

Habitat Variable	Basis for Determining Variable Value
Maximum channel depth	Determined from reach bathymetry
Off-channel bays	Calculated as 2008 average of stations in reach
Vertical wall banks	Measured using bank video, in conjunction with GIS
Riprap banks	Measured using bank video, in conjunction with GIS
Manmade structures	Determined from CAWS bank video
Macrophyte cover	Calculated as 2008 average of stations in reach
Percent overhanging vegetation	Length of riparian overhanging veg. for entire reach determined by inspection of bank video and recorded in GIS. Depth of overhang calculated as 2008 average measured at stations in each reach
Bank pocket areas	Calculated as 2008 average of stations in reach, validated using bank video
Large substrate – shallow	Calculated as 2008 average of stations in reach
Large substrate – deep	Calculated as 2008 average of stations in reach
Organic sludge	Calculated as 2008 average of stations in reach

Table 7-5: Values of Key Habitat Variables Assigned to Major CAWS Reaches.

Reach	Maximum Channel Depth (ft)	Off-Channel Bays	Vertical Wall Banks (%)	Riprap Banks (%)	Manmade Structures (average # per 400 m reach, 1 significant figure)	Macrophyte Cover (%)	Overhanging Vegetation (%)	Bank Pocket Areas	Large Substrate – Shallow (%)	Large Substrate – Deep (%)	Organic Sludge (%)
Upper North Shore Channel (North of North Side WRP)	8	2	0	0	1	9	33	0	20	0	0
Lower North Shore Channel	12	2	0	7	2	11	30	3	21	4	0
Upper North Branch Chicago River (North of Addison)	12	2	9	53	2	0	25	15	85	0	0
Lower North Branch Chicago River (South of Addison)	13	5	80	18	1	0	5	6	7	2	5
Chicago River	21	8	97	0	1	0	0	0	19	0	6
South Branch Chicago River	23	7	90	4	2	0	0	5	1	9	6
Bubbly Creek	13	1	35	3	2	0	8	9	5	5	48
Chicago Sanitary and Ship Canal	26	4	59	5	2	1	5	12	24	3	7
Cal-Sag Channel	16	2	19	53	2	0	5	12	24	5	9
Little Calumet River	15	6	5	17	1	1	6	17	5	26	4

Using the CAWS habitat index equation presented in Section 7.3 and the values presented in Table 7-5, the CAWS habitat index score for each major reach can be calculated. As mentioned in the preceding section, the raw values of the index were used for station-by-station scoring during index development, but for scoring of reaches and for other applications, the index is normalized to a scale of zero to 100.

The normalization process was performed by assigning probable worst case and best case values to each habitat variable and calculating the resulting index values. For variables that are unlikely to change in the CAWS, such as maximum depth, the existing range of values was used to establish the worst and best cases. For bank condition variables, a range of zero to 100% was used because these variables could possibly be altered beyond what presently exists at a given location in the CAWS. The worst case and best case values and the calculated index scores are presented in Table 7-6.

Table 7-6: Worst Case and Best Case Values Assigned to Habitat Variables for Normalization of CAWS Habitat Index.

Variable	Transformed Value		Transformed Value	
	Value	Worst Case	Value	Best Case
Constant:		12.8		12.8
MAX_DEP	26	9.91	6	0.38
OFF_CH_BAY	0	0	9	2.37
BNK_WALL	100	3.19	0	0.00
BNK_RIPRAP	100	5.12	0	0.00
MAN_MADE_STRUC	4	9.75	0	0
MCRPH_CHAN	0	0	13	2.78
PER_COV_ALT	0	0	33	3.3
BANK_POC_AREA	0	0	20	1
BIG_S	0	0	85	0.43
BIG_D	0	0	30	0.15
CAWS_ORGSLG	48	3.84	0	0
Raw CAWS Habitat Index:		-19.01		22.45
Final CAWS Habitat Index:		**0**		**100**

After assigning worst case and best case values to each variable, the values were transformed using the transformations shown in the regression equation and summed to obtain a RAW index score (-19.01 to 22.45). The final index value was calculated by adding the minimum score (19.01) to the raw index, dividing that by the range of raw values (22.45 – (-19.01) = 41.46), and multiplying by 100. The results are summarized in Table 7-7 and depicted in Figure 7-2.

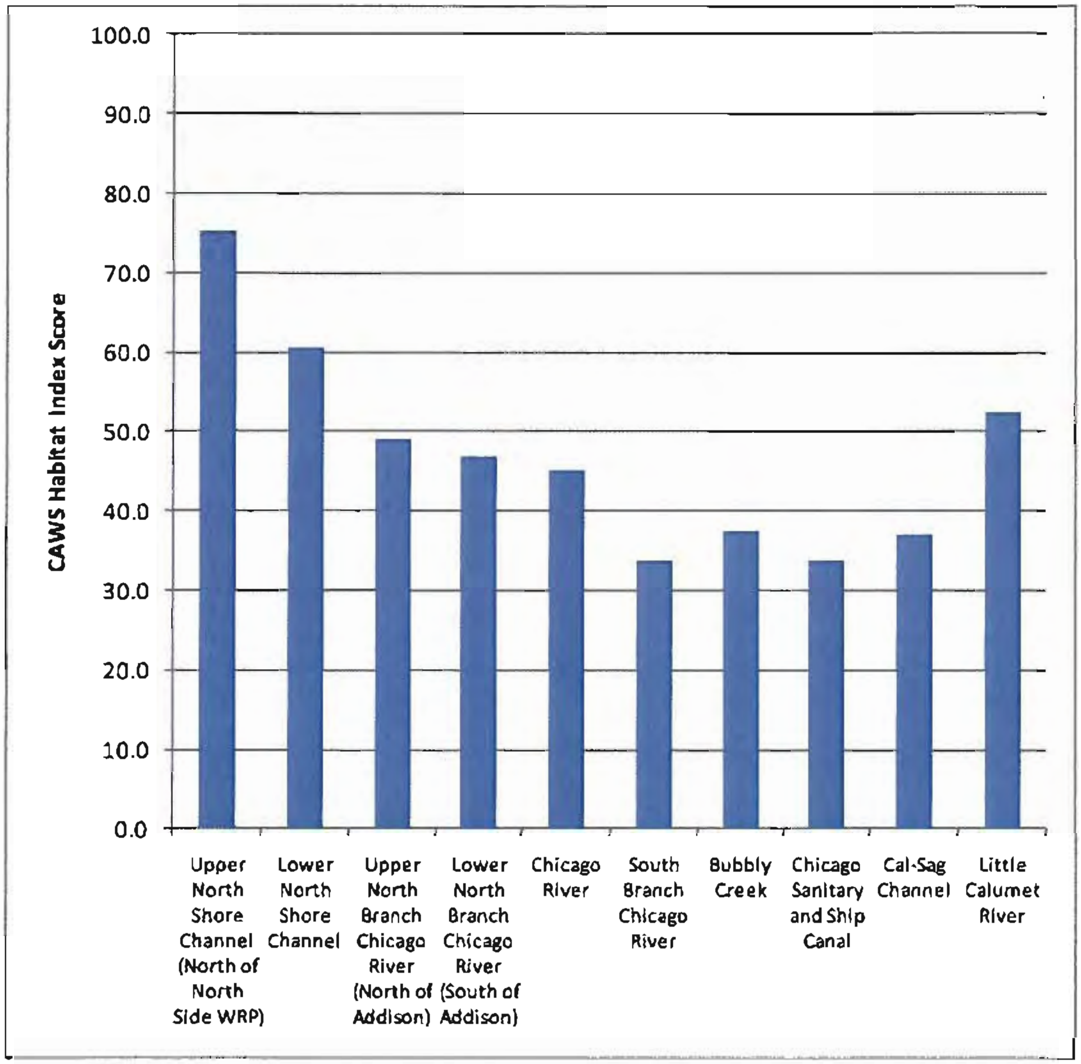

Figure 7-2: Results of CAWS Habitat Index Scoring for Major CAWS Reaches.

Table 7-7: CAWS Habitat Index Scores for Major Reaches.

Reach	CAWS Habitat Index Score
North Shore Channel North of North Side WRP	75.2
North Shore Channel South of North Side WRP	60.4
North Branch Chicago River North of Addison	49.1
North Branch Chicago River South of Addison	46.9
Chicago River	45.0
South Branch Chicago River	33.8
Bubbly Creek	37.4
Chicago Sanitary and Ship Canal	33.8
Cal-Sag Channel	37.1
Little Calumet River	52.4

7.5 POTENTIAL LIMITATION OF THE CAWS HABITAT INDEX

The CAWS Habitat Index (CHI) described above will provide a reasonable measure of physical habitat quality in the CAWS, to the extent that such a relationship can be developed with existing data. However, it is recognized that the data used to develop this index can be improved. Specifically, data were not available to adequately evaluate underwater habitat conditions in the CAWS, such as the presence of submerged structures. Because much of the system is maintained for navigation and effluent conveyance, large structures like fallen trees are routinely removed. Nonetheless, limited investigation during this Study using side scan sonar revealed the presence of some large woody debris and other submerged structures that might provide in-stream cover for fish. However, lacking sufficient data on submerged structure, it was not possible to evaluate its potential importance to fish in this Study. Further investigation of the potential for side scan sonar or some other remote sensing technology to observe and quantify the presence of submerged structure in the CAWS is recommended.

In spite of this limitation, the index presented here is useful in better understanding the relative differences in physical habitat in the CAWS.

This page is blank to facilitate double sided printing

8. SUMMARY OF CAWS HABITAT EVALUATION

The data and analyses described in the preceding sections were used to conduct a comprehensive evaluation of physical habitat in the CAWS. The evaluation documented in this Study is summarized, including major conclusions.

8.1 MAJOR CONCLUSIONS

Several major conclusions are supported by the work conducted in this study, including the following:

- ***Aquatic habitat is inherently limited in the CAWS by the system's form and function.*** Habitat in the CAWS is significantly limited by the design of the CAWS, most of which is manmade. The manmade reaches of the CAWS were built to support wastewater effluent conveyance and commercial navigation. The reaches that were once natural streams have been heavily modified to serve these purposes and the changes are unlikely to be reversed as long as the CAWS needs to serve these functions. The form and uses of the CAWS impose severe limitations on physical habitat in the system.

- ***Physical habitat is relatively more important to fish in the CAWS than dissolved oxygen.*** When key physical habitat variables and dissolved oxygen metrics are statistically compared to fish data collected between 2001 and 2008 in the CAWS, it is apparent that habitat is much more important to fish than dissolved oxygen. Multiple linear regression shows that the dominant habitat variables identified in this study had an r-squared of 0.48 with fish, indicating that these habitat variables explain as much as 48%, or about half, of the variability in the fish data.

- ***Explaining approximately half of the CAWS fish data variability is excellent, considering the natural variability in the fish data itself.*** As stated above, about half of the variability in fish data in the CAWS is explained by physical habitat, in particular certain key habitat variables identified in this study. Of the half of fish data variability not explained by the key habitat variables, most is explainable by natural variation in the fish data from one sampling event to another at each location. In other words, fish samples exhibit large temporal variability at any given location in the CAWS and when the portion of fish data variability not explained by habitat is statistically analyzed, it is most related to the variation at sampling locations over time, independent of habitat changes.

- ***Dissolved oxygen is relatively poor at explaining variability in fish data in the CAWS.*** Dissolved oxygen does not, for the most part, have a statistically significant relationship with fish in the CAWS. Various measures of dissolved oxygen were tested, including compliance with existing and proposed water quality standards, average and minimum DO, and percent of time below various DO concentration thresholds. The strongest relationship identified between any of these metrics and the fish data had an r-squared value of 0.27, which is about half as good as the key habitat variables identified in this study. All other DO measures tested had r-squared values significantly lower than this. This indicates that physical habitat, not water quality, is the most limiting factor for fish in the CAWS today.

Some further elaboration on these conclusions is provided in the sections below.

8.2 SUMMARY OF KEY HABITAT VARIABLES

The process described in Sections 6 and 7 of this report used fish and habitat data collected from throughout the CAWS to identify the physical habitat variables most closely correlated with fish metrics in the CAWS. Those variables are:

- Maximum depth of channel
- Off-channel bays
- Percent of vertical wall banks in reach
- Percent of riprap banks in reach
- Manmade structures in reach
- Percent macrophyte cover in reach

Many of these key habitat variables are the result of the major functions that the CAWS serves. Channel depth, vertical wall banks, and riprap are all the result of the need to support commercial navigation, effluent conveyance, flood control, or all three. Other habitat variables are so uniformly absent or of such uniformly poor quality in the CAWS as a result of the origin, design and function of the CAWS that they do not register as important. These include habitat attributes that are normally important in natural systems such as substrate, in-stream cover, floodplain connectivity, and morphological variation.

Using multiple linear regression analyses, the key habitat variables listed above were able to explain 48% of the variability in fish data collected from the CAWS from 2001 – 2007. Additional analyses described in Section 7.5.2 show that most of the variability in the 2008 fish data not explained by these physical habitat variables was attributable to variability in the fish sampling results. DO was also shown to be relatively less important in explaining fish data variability than these key habitat variables.

8.3 RELATIVE IMPORTANCE OF PHYSICAL HABITAT IN THE CAWS

As stated above, the regression analysis presented in Section 6.3.3 shows that physical habitat alone can explain up to 48% of the variance in fish data collected in the CAWS from 2001 – 2007, which is significantly better than can be accomplished by evaluating water quality alone. In the analysis presented in Appendix C, the DO metric most highly correlated with fish data only had an r-squared of 0.27, meaning that DO alone can only explain 27% of the variability in the same seven years of fish data. Other important findings include:

- Of the 52% of fish data variability that is not explained by these physical habitat variables, as much as 70% of it can be explained by variation in fish sampling results from year to year. This means that the fish measured at a location can vary significantly from one sample event to the next and that this will lead to an inherent variability in the data that cannot be explained by changes in independent variables such as habitat or water quality.

- The percent of time that DO is less than 5 mg/L at a given station in the CAWS from June through September, which was the water quality metric most closely correlated with fish, explains approximately 3% of the 52% of the fish data variability that is not explained by the six key physical habitat variables.

- When the key DO metric is included with the six key habitat variables in the regression with fish data, the ability of the regression to explain variability in fish data is only increased by 4% over using habitat alone.

All of these findings indicate that physical habitat is relatively more important than water quality to fish in the CAWS.

8.4 OTHER RELEVANT HABITAT CONSIDERATIONS

It should be noted that, while the analysis conducted in this study led to the identification of key habitat variables, it is very much a data-driven analysis and although two separate data sets were used for the quantification of the relationship between habitat and fish, and the testing of that relationship, there are almost certainly other habitat factors that are or could be of value to aquatic life in the CAWS. These may include the following:

- Submerged structure: As discussed elsewhere in this report, no complete data on submerged structure were collected in this Study, although pilot testing of side s can sonar indicates that there may be value in using that technology to image subsurface conditions and identify submerged structure. If submerged structure can be quantified and if there is sufficient submerged structure in the CAWS to support statistical analysis, it may be possible to identify a relationship between submerged structure and fish in the CAWS.

- Off-channel habitat: Because of the channelized, constructed, and urban nature of the CAWS, there is little connected, off-channel habitat. Such areas can provide habitat for different life stages of fish as well as refuge. In the CAWS, they may provide shelter from boat wakes. In the general absence of such features, it is not possible to evaluate their potential value to aquatic life in the CAWS at present, because insufficient data exist.

- Navigation: Although there are insufficient data at present to quantify the specific effects of navigation on fish in the CAWS, the impacts almost certainly are occurring and cannot be ignored. Further research would be required to document and quantify these impacts, but navigation clearly presents significant limitations to aquatic biota in the CAWS, both through limitations imposed on physical habitat and through direct effects. The channel design/modification to support navigation presents significant limitations to the habitat improvement potential in the CAWS.

While these and other aspects of physical habitat are not represented in the CAWS habitat index, it does not mean that they are not important, it simply means that they either are not present in sufficient quantity within the CAWS or have not been fully measured to date.

9. REFERENCES

Allan, J.D. 1995. *Stream Ecology: Structure and Function of Running Waters.* Dordrecht, The Netherlands: Kluwer Academic Publishers. 1995.

Alp, E. and Melching, C.S., 2008. *Calibration of a Model for Simulation of Water Quality During Unsteady Flow in the Chicago Waterway System and Application to Evaluate Use Attainability Analysis Remedial Actions.* Published as Metropolitan Water Reclamation District of Greater Chicago Research and Development Department Report No. 06-84. February 2006.

Arlinghaus, R., Englehardt, C., Sukhodolov, A., and Wolter, C., 2002. *Fish recruitment in a canal with intensive navigation: implications for ecosystem management.* Journal of Fish Biology. Vol. 61, pp1386-1402. 2002.

Armantrout, N. B., 1998. "Glossary of Aquatic Habitat Inventory Terminology." American Fisheries Society. Bethesda, Maryland.

Bain, M. B. and Stevenson, N. J., eds., 1999. *Aquatic Habitat Assessment: Common Methods*, American Fisheries Society. Bethesda, Maryland.

Barbour, M. T., Paul, M. J., Bressler, D. W., O'Dowd, A. P., Resh, V. H., and Rankin, E., 2007. "Bioassessment: A Tool for Managing Aquatic Life Uses for Urban Streams." Water Environment Research Foundation. Alexandria, Virginia.

Barlaup, B. T., Gabrielsen, S. E., Skoglund, H., and Wiers, T., 2008. "Addition of Spawning Gravel-A Means to Restore Spawning Habitat of Atlantic Salmon (Salmo Salar L.), and Anadromous and Resident Brown Trout (Salmo Trutta L.) in Regulated Rivers," *River Research and Applications*, Vol. 24, pp. 543-550.

Becker, G. C., 1983. *Fishes of Wisconsin*, The University of Wisconsin Press. Madison, Wisconsin.

Blocksom, K.A. and Flotemersch, J.E. 2005. *Comparison of Macroinvertebrate Sampling Methods for Nonwadeable Streams.* Environmental Monitoring and Assessment. Vol 102, pp243-262. 2005.

Blocksom, K. A., and Flotermersch, J. E., 2008. "Field and Laboratory Performance Characteristics of a New Protocol for Sampling Riverine Macroinvertebrate Assemblages," *River Research and Applications*, Vol. 24, pp. 373-387.

Booth, D. B., Hartley, D., and Jackson, R., 2002. "Forest Cover, Impervious-Surface Area, and the Mitigation of Stormwater Impacts," *Journal of the American Water Resources Association*, Vol. 38, No.3., pp. 835-845.

Bradley, D. L., 2008. "Field Work Correspondence With Anglers During 2008 Field Work Within the Chicago Area Waterway System. Doug Bradley (LimnoTech, Inc.,) June - September 2008."

Brammeier, J., Polls, I., and Mackey, S., 2008. *Preliminary Feasibility of Ecological Separation of the Mississippi River and the Great Lakes to Prevent the Transfer of Aquatic Invasive Species.* Great Lakes Fishery Commission 2008 Project Completion Report.

Brown, L. R., Gray, R. H., Hughes, R. M., Meador, M. R., eds., 2005. *Effects of Urbanization on Stream Ecosystems*, American Fisheries Society, Symposium 47. Bethesda, Maryland.

Bunn, S. E., and Arthington, A. H., 2002. "Basic Principles and Ecological Consequences of Altered Flow Regimes for Aquatic Biodiversity," *Environmental Management*, Vol. 30, No.4., pp. 492-507.

Burton, G. A., and Landrum, P. F., 2003. "Toxicity of Sediments," in the *Encyclopedia of Sediments and Sedimentary Rocks*. The Netherlands: Kluwer Academic Publishers. pp. 748-751.

CDM, 2007. "Chicago Area Waterway System Use Attainability Analysis Final Report." Illinois Environmental Protection Agency.

Coles, J. F., Cuffney, T. F., McMahon, G., and Beaulieu, K. M., 2004. "The Effects of Urbanization on the Biological, Physical, and Chemical Characteristics of Coastal New England Streams." Professional Paper No. 1695, U.S. Geological Survey.

D'Ambrosio, J. L., Williams, L. R., Witter, J. D., and Ward, A., 2009. "Effects of Geomorphology, Habitat, and Spatial Location on Fish Assemblages in a Watershed in Ohio, USA," *Environmental Monitoring and Assessment*, Vol. 148, pp. 325-341.

Draper, N. R. and H. Smith, 1981. *Applied Regression Analysis*, John Wiley & Sons. New York, NY.

Duffy-Anderson, J.T., Manderson, J.P., and Able, K.W., 2003. *A Characterization of Juvenile Fish Assemblages Around Man-Made Structures in the New York-New Jersey Harbor Estuary, U.S.A.* Bulletin of Marine Science. Vol. 72, No. 3, pp877-899. 2003.

Edwards, C.J., Griswold, B.L., Tubb, R.A., Weber, E.C., and Woods, L.C., 1984. *Mitigating Effects of Artificial Riffles and Pools on the Fauna of a Channelized Warmwater Stream.* North American Journal of Fisheries Management. Vol. 4, pp194-203. 1984.

Emery, E. B., Simon, T. P., McCormick, F. H., Angermeier, P. L., Deshon, J. E., Yoder, C. O., Sanders, R. E., Pearson, W. D., Hickman, G. D., Reash, R. J., and Thomas, J. A., 2003. "Development of a Multimetric Index for Assessing the Biological Condition of the Ohio River," *Transactions of the American Fisheries Society*, Vol. 132, pp. 791-808.

Fischenich, J. Craig, 2003. *Effects of Riprap on Riverine and Riparian Ecosystems.* Washington, DC, U.S. Army Corps of Engineers.

Fitzpatrick, F. A., Waite, I. R., D'Arconte, P. J., Meador, M. R., Maupin, M. A., and Gurtz, M. E., 1998. *Revised Methods for Characterizing Stream Habitat in the National Water-Quality Assessment Program.* US Geological Survey Water-Resources Investigations Report 98-4052.

Flotemersch, J. E., and Blocksom, K. A., 2005. "Electrofishing in Boatable Rivers: Does Sampling Design Affect Bioassessment Metrics?" *Environmental Monitoring and Assessment*, Vol. 102, pp. 263-283.

Flotemersch, J.E., Stribling, J.B., and Paul, M.J., 2006. *Concepts and Approaches for the Bioassessment of Non-Wadeable Streams and Rivers.* EPA 600-R-06-127. Cincinnati, Ohio: USEPA National Exposure Research Laboratory.

Gabel, F., Garcia, X.-F., Brauns, M., Suckhodolov, A., Leszinski, M., and Pusch, M. T., 2008. "Resistance to Ship-Induced Waves of Benthic Invertebrates in Various Littoral Habitats," *Freshwater Biology*, Vol. 53, pp. 1567-1578.

Giller, P. S., and Malmquist, B., eds., 1998. *The Biology of Streams and Rivers*, Oxford University Press.

Gordon, N. D., McMahon, T. A., Finlayson, B. L., Gippel, C. J., and Nathan, R. J., eds., 2004. *Stream Hydrology: An Introduction for Ecologists*, Second Edition ed., John Wiley & Sons, Ltd.

Greenberg, J., 2002. *A Natural History of the Chicago Region.* The University of Chicago Press.

Gutreuter, S., Dettmers, J. M., and Wahl, D. H., 2003. "Estimating Mortality Rates of Adult Fish From Entrainment Through the Propellers of River Towboats," Transactions of the American Fisheries Society, Vol. 132, pp. 646-661.

Hall, L.W., Morgan, R.P., Perry, E.S., and Waltz, A., 1999. *Development of a Provisional Physical Habitat Index for Maryland Freshwater Streams.* Annapolis, MD: Maryland Department of Natural Resources. 1999.

Hill, Libby. 2000. *The Chicago River: A Natural and Unnatural History.* Chicago: Lake Claremont Press. 2000.

Holland, L. E., 1987. "Effect of Brief Navigation-Related Dewaterings on Fish Eggs and Larvae," North American Journal of Fisheries Management, Vol. 7, pp. 145-147.

Illinois Department of Natural Resources (IDNR), 2000. "Draft Manual for Calculating Index of Biotic Integrity Scores for Streams in Illinois."

Karr, J. R., 1981. "Assessment of Biotic Integrity Using Fish Communities," *Fisheries*, Vol. 6, No.6., pp. 21-27.

Karr, J. R., 1995. "Protecting Aquatic Ecosystems: Clean Water Is Not Enough," *Biological Assessment and Criteria: Tools for Water Resource Planning and Decision Making*. W. S. Davis, and T. P. Simon, eds., Lewis Publishers, pp. 7-13.

Karr, J. R., and Yoder, C. O., 2004. "Biological Assessment and Criteria Improve Total Maximum Daily Load Decision Making," *Journal of Environmental Engineering*, Vol. 130, No.6., pp. 594-604.

Kaufmann, P. R., 2000. "Physical Habitat Characterization - Non-Wadeable Rivers," *Environmental Monitoring and Assessment Program - Surface Waters: Field Operations and Methods for Measuring the Ecological Condition of Non-Wadeable Rivers and Streams*. J. M. Lazorchak, B. H. Hill, D. K. Averill, D. V. Peck, and D. J. Klemm, eds., US Environmental Protection Agency, Cincinnati, Ohio.

Kennen, J. G., Chang, M., and Tracy, B. H., 2005. "Effects of Landscape Change on Fish Assemblage Structure in a Rapidly Growing Metropolitan Area in North Carolina, USA," *Effects of Urbanization on Stream Ecosystems*. L. R. Brown, R. H. Gray, R. H. Hughes, and M. R. Meador, eds., American Fisheries Society, Symposium 47, Bethesda, Maryland. pp. 39-52.

Kohler, C.C. and Hubert, W.A., eds. 1999. *Inland Fisheries Management in North America, Second Edition*. American Fisheries Society, Bethesda, Maryland.

Lyons, J., Piette, R. R., and Niermeyer, K. W., 2001. "Development, Validation, and Application of a Fish-Based Index of Biotic Integrity for Wisconsin's Large Warmwater Rivers," *Transactions of the American Fisheries Society*, Vol. 130, pp. 1077-1094.

MacDonald, D. D., and Ingersol, C. G., 2002. "A Guidance Manual to Support the Assessment of Contaminated Sediments in Freshwater Ecosystems, Volume 1." Rep. No. EPA-905-B02-001-A, USEPA.

McMahon, T. E., Zale, A. V., and Orth, D. J., 1996. "Aquatic Habitat Measurements," *Fisheries Techniques*. B. R. Murphy, and D. W. Willis, eds., American Fisheries Society, Bethesda, Maryland. pp. 83-120.

Merritt, R. W., Allan, J. D., Cummins, K. W., Wessell, K. J., and Wilhelm, J. O., 2005. "Qualitative Biological and Habitat Protocols for Michigan's Non-Wadeable Rivers." Michigan Department of Environmental Quality. Lansing, Michigan.

Metropolitan Water Reclamation District of Greater Chicago (District). *Ambient Water Quality Monitoring Quality Assurance Project Plan, Revision 2.1.* November 1, 2007.

Metropolitan Water Reclamation District of Greater Chicago (District) 2008. *Description of the Chicago Waterway System for the Use Attainability Analysis*. District Research and Development Department Report No. 08-15. March 2008.

Minarik, Thomas. E-mail correspondence to LimnoTech dated September 4, 2009.

Minarik, T.A., J.L. Wasik, M. Sopcak, and S.G. Dennison. *Continuous Dissolved Oxygen Monitoring in the Deep Draft Chicago Waterway System During 2007*. MWRDGC Research and Development Department. August 2008.

Morgan, R. P. I., Ulanowicz, R. E., Rasin, V. J. Jr., Noe, L. A., and Gray, G. B., 1976. "Effects of Shear on Eggs and Larvae of Striped Bass, Morone Saxatilis, and White Perch, M. Americana," Transactions of the American Fisheries Society, Vol. 1, pp. 149-154.

Ohio Environmental Protection Agency (OEPA), 1989. "Biological Criteria for the Protection of Aquatic Life. Vol. III. Standardized Field Sampling and Laboratory Methods for Assessing Fish Sampling and Macroinvertebrate Communities." Ohio Environmental Protection Agency, Division of Water Quality Monitoring and Assessment. Columbus, Ohio.

Orth, D.J. and R.J. White 1999. "Stream Habitat Management" *Chapter 10 of Inland Fisheries Management in North America (2nd Ed.).* Bethesda, MD: American Fisheries Society, 1999.

Penczak, T., O'Hara, K., and Kostrzewa, J., 2002. "Fish Bioenergetics Model Used for Estimation of Food Consumption in a Navigation Canal With Heavy Traffic," Hydrobiologia, Vol. 479, pp. 109-123.

Pott, D. B., 2009. "Technical Memorandum No. 1: Characterization of the Macroinvertebrate Community." Baetis Environmental Services, Inc. Chicago, Illinois.

Rabeni, C. F., and Jacobson, R. B., 1999. "Warmwater Streams," *Inland Fisheries Management in North America Second Edition*. C. C. Kohler, and W. A. Hubert, eds., American Fisheries Society, Bethesda, Maryland. pp. 505-528.

Ramey, H.P., 1953. *Diversion of Water From Lake Michigan*. No publisher information available.

Rankin, E. T., 1989. "The Qualitative Habitat Evaluation Index (QHEI): Rationale, Methods and Application." Ohio Environmental Protection Agency, Division of Water Quality Planning and Assessment. Columbus, Ohio.

Rankin, E. T., 1995. "Chapter 13: Habitat Indices in Water Resource Quality Assessments," *Biological Assessment and Criteria Tools for Water Resource Planning and Decision Making*. W. S. Davis, and T. P. Simon, eds., Lewis Publishers, pp. 181-208.

Rankin, E. T., 2004. "Analysis of Physical Habitat Quality and Limitations to Waterways in the Chicago Area."

Reash, R. J., 1999. "Considerations for Characterizing Midwestern Large-River Habitats," *Assessing the Sustainability and Biological Integrity of Water Resources Using Fish Communities*. Thomas P.Simon, ed., CRC Press, pp. 463-473.

Resh, V.H., A.V. Brown, A.P. Covich, M.E. Gurtz, H.W.Li, G.W. Minshall, S.R. Rice, A.L. Sheldon, J.B. Wallace, and R.C. Wissmar 1988. *The Role of Disturbance in Stream Ecology*. Journal of the North American Benthological Society. Vol. 7, No. 4, pp433-455. 1988.

Roset, N., Grenouillet, G., Goffaux, D., Pont, D., and Kestemont, P., 2007. "A Review of Existing Fish Assemblage Indicators and Methodologies," *Fisheries Management and Ecology*, Vol. 14, pp. 393-405.

Rosgen, D. 1996. *Applied River Morphology*. Pagosa Springs, CO: Wildland Hydrology, 1996.

Schramm, H. Jr., Minnis, R. B., Spencer, A. B., and Theel, R. T., 2008. "Aquatic Habitat Change in the Arkansas River After the Development of a Lock-and-Dam Commercial Navigation System," River Research and Applications, Vol. 24, pp. 237-248.

Sheehan, R. J., and Rasmussen, J. L., 1999. "Large Rivers," *Inland Fisheries Management in North America*. Christopher C.Kohler, and Wayne A.Hubert, eds., American Fisheries Society, Bethesda, Maryland.

Short, T.M., E.M. Giddings, H. Zappia, and J.F. Coles 2005. *Urbanization Effects on Stream Habitat Characteristics in Boston, Massachusetts; Birmingham, Alabama; and Salt Lake City, Utah*. American Fisheries Society Symposium. Vol. 47, pp317-332. 2005.

Simon, T. P., and Lyons, J., 1995. "Application of the Index of Biotic Integrity to Evaluate Water Resource Integrity in Freshwater Ecosystems," *Biological Assessment and Criteria: Tools for Water Resource Planning and Decision Making*. W. S. Davis, and T. P. Simon, eds., Lewis Publishers, pp. 245-262.

Simon, T. P., and Sanders, R. E., 1999. "Applying an IBI Based on Great-River Fish Communities," *Assessing the Sustainability and Biological Integrity of Water Resources Using Fish Communities*. T.P.Simon, ed., CRC Press, Boca Raton, Florida. pp. 475-505.

Solzman, D.M. 2006. *The Chicago River: An Illustrated History and Guide to the River and its Waterways, 2nd edition.* Chicago: University of Chicago Press.

Smogor, R.A. and P.L. Angermeier 1999. "Relations Between Fish Metrics and Measures of Anthropogenic Disturbance in Three IBI Regions in Virginia." Chapter 23 of *Assessing the Sustainability and Biological Integrity of Water Resources Using Fish Communities*. Boca Raton, FL: CRC Press.

Tate, C., T.F. Cufney, G. MacMahon, E.P. Giddings, J.F. Coles, and H. Zappia 2005. *Use of an Urban Intensity Index to Assess Urban Effects on Streams in Three Contrasting Environmental Settings*. American Fisheries Society Symposium. Vol. 47, pp291-315. 2005.

United States Environmental Protection Agency (USEPA), 2005. "Draft Use of Biological Information to Better Define Designated Aquatic Life Uses in State and Tribal Water Quality Standards: Tiered Aquatic Life Uses."

United States Environmental Protection Agency (USEPA)."Contaminated Sediment in Water: Basic Information," May 18, 2008. Accessed online at http://www.epa.gov/waterscience/cs/aboutcs/ on Mar. 19, 2009.

Waters, T. F. 1995. *Sediment in Streams: Sources, Biological Effects, and Control*, American Fisheries Society Monograph 7. Bethesda, Maryland.

Weigel, B. M., Lyons, J., and Rasmussen, P. W., 2006. "Fish Assemblages and Biotic Integrity of a Highly Modified Floodplain River, The Upper Mississippi, and a Large, Relatively Unimpacted Tributary, The Lower Wisconsin," *River Research and Applications*, Vol. 22, pp. 923-936.

Wesche, T. A., and Isaak, D. J., 1999. "Watershed Management and Land Use Practices," *Inland Fisheries Management in North America, 2nd Edition*. C. C. Kohler, and W. A. Hubert, eds., American Fisheries Society, Bethesda, Maryland. pp. 217-248.

Wilhelm, J. G. O. 2002. "A Habitat Rating System for Non-Wadeable Rivers of Michigan." Master of Science, University of Michigan.

Wilhelm, J.G.O., J.D. Allan, K.J. Wessell, R.M. Merritt, and K.W. Cummins 2005. *Habitat Assessment of Non-Wadeable Rivers in Michigan.* Environmental Management. Vol. 36, No. 4, pp592-609. 2005.

Wolter, C., 2001. "Conservation of Fish Species Diversity in Navigable Waterways," Landscape and Urban Planning, Vol. 53, pp. 135-144.

Wolter, C. and R. Arlinghaus 2003. *Navigation impacts on freshwater fish assemblages: the ecological relevance of swimming performance.* Reviews in Fish Biology and Fisheries. Vol. 13, pp63-89. 2003.

Wolter, C., Arlinghaus, R., Sukhodolov, A., and Engelhardt, C., 2004. "A Model of Navigation-Induced Currents in Inland Waterways and Implications for Juvenile Fish Displacement," Environmental Management, Vol. 34, No.5., pp. 656-668.

Wysocki, L. E., Dittami, J. P., and Ladich, F., 2006. "Ship Noise and Cortisol Secretion in European Freshwater Fishes," Biological Conservation, Vol. 128, pp. 501-508.

APPENDIX A:

REPORT ON FISH METRIC SELECTION FOR THE CAWS HABITAT EVALUATION AND IMPROVEMENT STUDY

This page is blank to facilitate double sided printing