**Indices Information**

**HGMI (Genus)**

The HGMI genus classification is based on seven metrics: Total number of Genera, % Not Insect Taxa, % Sensitive EPT, Hilsenhoff Biotic Index (HBI), # of Scraper Genera, Attribute 2 Genera, and Attribute 3 Genera. Four (Total number of Genera, % Sensitive EPT, HBI, # of Scraper Genera) are adjusted for watershed site using the formulas below.

|  |  |
| --- | --- |
| Total Number of Genera | =METRIC+26.53(22.776+4.173\*(log (area\_ sq\_km))) |
| % Not Insect Taxa | =METRIC |
| % Sensitive EPT | = (METRIC)+37.49-(49.22 -13.8\*(log (area\_sq\_km))) |
| HBI | =METRIC + 4.23-(3.407+0.981\*(log(area\_sq\_km))) |
| # of Scraper Genera | =METRIC + 5.44-(3.889+0.981\*(log (area\_sq\_km))) |
| Attribute 2 Genera | =METRIC |
| Attribute 3 Genera | =METRIC |

In this case percentages refer to whole numbers (ex. For 12%, the metric will be 12 not 0.12)

These values are then corrected using the scoring formulas below.

|  |  |
| --- | --- |
| **Scoring Formulas** | |
| Total Number of Genera | =100\*METRIC/31 |
| % Not Insect Taxa | =100\*(60-METRIC)/55 |
| % Sensitive EPT | =100\*METRIC/69 |
| HBI | =100\*(7.2-METRIC)/4.6 |
| # Scraper Genera | =100\*METRIC/11 |
| Attribute 2 Genera | =100\*METRIC/8 |
| Attribute 3 Genera | =100\*METRIC/8 |

The HGMI score can be found be adding together the values from the scoring formulas and dividing by seven.

The water quality can then be classified based on the final score.

|  |  |
| --- | --- |
| <\_\_ | Excellent |
|  | Good |
|  | Fair |
|  | Poor |

**HGMI (Family)**

The HGMI Genus classification is based on five metrices: Number of Families, % Not Insect Taxa, % Sensitive EPT, Family Biotic Index (FBI), and # of Scraper Genera. Two (% Sensitive EPT, FBI) are adjusted for watershed site using the formulas below

|  |  |
| --- | --- |
| Number of families | =METRIC |
| % Not Insect Families | =METRIC |
| % Sensitive EPT | =METRIC+35.15-(45.59-11.59\*(log(area\_sq\_km))) |
| FBI | =METRIC+4.19-(3.636+0.615\*(log(area\_sq\_km))) |
| # of Scraper Genera | =METRIC |

These values are then corrected using the scoring formulas below.

|  |  |
| --- | --- |
| **Scoring Formulas** | |
| Number of families | =100\*METRIC/12 |
| % Not Insect Families | =100\*(70-(METRIC))/63 |
| % Sensitive EPT | =100\*METRIC/67 |
| FBI | =100\*(7-METRIC)/4 |
| # of Scraper Genera | =100\*METRIC/6 |

The HGMI score can be found be adding together the values from the scoring formulas and dividing by seven.

The water quality can then be classified based on the final score.

**CPMI (genus)**

\_\_ is based on \_\_\_\_

**NJIS**

**Sheets**

|  |  |
| --- | --- |
| **Sheet** | **Contents** |
| SITE INFORMATION DATA | Information about site, data, what sheet the data is stored in, and its original location |
| METADATA | Information about any edits made to data from the original sheet |
| (SHEETS WITH SITE CODES) - RED | Calculations and scores for the NJIS, HGMI (genus), and HGMI (family) |
| (SHEETS WITH SITE CODES) - BLUE | Calculations and scores for the NJIS and CPMI |
| DATA\_WITH\_GENUS\_NEW | All raw data that included information about the genus, sample of 100, and information needed to calculate the indices. Data is from 2022 – 2024. |
| DATA\_WITH\_GENUS\_OLD | All raw data that included information about the genus, sample of 100, and information needed to calculate the indices. Data is from 2014 – 2017 and 2021. |
| DATA\_WITHOUT\_GENUS | All raw data without information about the genus and information needed to calculate the indices. Data is from 1996 – 2014 and 2018 – 2019. |
| BugList | Reference for information about species |
| SiteSize | Lists the area of in square kilometers for all high gradient sites |

**Important Formulas**

Any formula values highlighted in orange will change depending on row or samples

Any formula values highlighted in purple will change depending on the range

From Sites Tabs

The formulas will adjust depending on whether the data comes from the DATA\_WITH\_GENUS\_NEW, DATA\_WITH\_GENUS\_OLD or DATA\_WITHOUT\_GENUS Sheet

Most formulas remain the same and just changing out the sheet name. The formulas to calculate the indices (headers highlighted grey) remain the same for all values. More information is given below.

For DATA\_WITH\_GENUS\_NEW

|  |  |  |
| --- | --- | --- |
| **Formula** | **Purpose** | **Location** |
| COUNTA(UNIQUE(FILTER(DATA\_WITH\_GENUS\_NEW!A:A&DATA\_  WITH\_GENUS\_NEW!C:C,(DATA\_  WITH\_GENUS\_NEW!A:A=A[ROW])))) | Counts number of unique families with the same site code | Column G |
| UNIQUE(FILTER(DATA\_WITH\_  GENUS\_NEW!C:C,(DATA\_WITH\_  GENUS\_NEW!A:A=A[ROW])\*(DATA\_WITH\_GENUS\_NEW!J:J=J[ROW])\*ROW(DATA\_WITH\_GENUS\_NEW!C:C),0)) | Determines the dominant family or families | Column K |

For DATA\_WITH\_GENUS\_OLD

|  |  |  |
| --- | --- | --- |
| **Formula** | **Purpose** | **Location** |
| COUNTA(UNIQUE(FILTER(  DATA\_WITH\_GENUS!A:A&  DATA\_WITH\_GENUS\_OLD!D:D,(DATA\_WITH\_GENUS\_OLD!A:A=A[ROW])\*(DATA\_WITH\_GENUS!J:J>0)))) | Counts number of unique families with the same site code | Column G |
|  |  |  |

For DATA\_WITHOUT GENUS

|  |  |  |
| --- | --- | --- |
| **Formula** | **Purpose** | **Location** |
| COUNTIFS(DATA\_WITHOUT\_  GENUS!A:A,A[ROW],DATA\_WITHOUT\_GENUS!I:I, ">0") | Counts number of unique families with the same site code | Column H |
|  |  |  |

From DATA\_WITH\_GENUS\_NEW

Data comes from the Normandeau 2022 to 2024 spreadsheets and from 2018 to 2019 data sheets. FTV, Tolerance value, Scraper classification, and Talu attribute were entered manually. Insect classification was imported from the BugList sheet based on family name. EPT classification was determined from Order information.

100 sample

These formulas will need to be manually adjusted to fit the range of the sample

|  |  |  |
| --- | --- | --- |
| **Formula** | **Purpose** | **Location** |
| (F[ROW]/(SUMIFS(E:E,A:A,  A[ROW]))\*100) | Finds the percentage of the count is of the total | Column G |
| =INT(G[ROW]) | Finds the integer value of percentage \_\_\_ | Column H |
| =G[ROW]-H[ROW] | Finds the difference between the percentage value and integer value | Column I |
| =INT(G[ROW])+IF(100-SUM(INT(G$[BEGINNING\_OF\_ RANGE]:G$[END\_OF\_RANGE]))>=RANK(I[ROW],I$[BEGINNING\_OF\_RANGE]:I$[END\_OF\_  RANGE])+COUNTIF(  I$[BEGINNING\_OF\_RANGE]:  I[ROW],I[ROW])-1,1,0) | Decides whether to round the int up to make sure the sample reaches 100 specimens | Column J |

Dominant Taxon

\_\_ are calculated separately to \_\_\_ genus \_\_ from the same family

|  |  |  |
| --- | --- | --- |
| **Formula** | **Purpose** | **Location** |
| =IF(COUNTIF(C$[BEGINNING\_  OF\_RANGE]:C$[END\_OF\_  RANGE],C[ROW])>1,SUMIF(  C$[BEGINNING\_OF\_RANGE]:  C$[END\_OF\_RANGE],C[ROW],  I$[BEGINNING\_OF\_RANGE]:  I$[END\_OF\_RANGE),I[ROW]) | Finds the number of specimens for all families (taking in account listings with different genera) | Column J |
|  |  |  |

From DATA\_WITH\_GENUS\_OLD

Data was taken tblSampleDates ranging from 2014 to 2017 and 2021. Information about EPT classification, FTV, Tolerance value, Insect classification, Clinger classification, Scraper classification, and Talu attribute are all from BugList sheet based on genus name.

100 sample

These formulas will need to be manually adjusted to fit the range of the sample

|  |  |  |
| --- | --- | --- |
| **Formula** | **Purpose** | **Location** |
| (F[ROW]/(SUMIFS(E:E,A:A,  A[ROW]))\*100) | Finds the percentage of the count is of the total | Column G |
| =INT(G[ROW]) | Finds the integer value of \_\_ | Column H |
| =G[ROW]-H[ROW] | Finds the difference between the percentage value and integer value | Column I |
| =INT(G[ROW])+IF(100-SUM(INT(G$[BEGINNING\_OF\_ RANGE]:G$[END\_OF\_RANGE]))>=RANK(I[ROW],I$[BEGINNING\_OF\_RANGE]:I$[END\_OF\_  RANGE])+COUNTIF(  I$[BEGINNING\_OF\_RANGE]:  I[ROW],I[ROW])-1,1,0) | Decides whether to round the int up to make sure the sample reaches 100 specimens | Column J |

From DATA\_WITHOUT\_GENUS

Data comes from the BATSITES COLLECTED spreadsheet from dates 1996 to 2014 (if not included in tblSampleDate) and 2018 to 2019 data sheets. FTV, Insect classification, and Scraper classification are all from the BugList sheet based on family name. EPT classification was determined from Order information.

100 sample

These formulas will need to be manually adjusted to \_\_\_\_

|  |  |  |
| --- | --- | --- |
| **Formula** | **Purpose** | **Location** |
| (E[ROW]/(SUMIFS(E:E,A:A,  A[ROW]))\*100) | Finds the percentage of the count is of the total | Column F |
| =INT(F[ROW]) | Finds the integer value of \_\_ | Column G |
| =F[ROW]-G[ROW] | Finds the difference between the percentage value and integer value | Column H |
| =INT(F[ROW])+IF(100-SUM(INT(F$[BEGINNING\_OF\_ RANGE]:F$[END\_OF\_RANGE]))>=RANK(H[ROW],H$[BEGINNING\_OF\_RANGE]:H$[END\_OF\_  RANGE])+COUNTIF(H$[BEGINNING\_OF\_RANGE]:H[ROW]  ,H[ROW])-1,1,0) | Decides whether to round the int up to make sure the sample reaches 100 specimens | Column I |

If these formulas do not result in a sample of 100 and formula is correct, manually adjust values based on largest percent different and order of values. Any changed values should be highlighted in orange.

**TO BE FIXED**

* EPT Taxa Richness (will still count duplicates)

**NEXT STEPS**

* Consolidate raw data into one sheet
* Making formulas less reliant on manually entry to define a range
* Consolidate all high gradient sites into one sheet
* Consolidate all low gradient sites into one sheet
* Determine a better way to enter the data from \_\_\_\_