Appendix B

B1 Candidate Metrics

B2 Summary of input metrics in existing low gradient IBIs

B3 Non-target taxa

B4 Exclusion criteria for redundant taxa

B1 Candidate Metrics

List of candidate macroinvertebrate metrics that were calculated with the BioMonTools R package (<https://github.com/leppott/BioMonTools>).

|  |  |  |
| --- | --- | --- |
| Metric Name | Category | Description |
| nt\_total | RICH | number of taxa - total |
| nt\_Amph | RICH | number of taxa - Order Amphipoda |
| pi\_Amph | COMP | percent individuals - Order Amphipoda |
| pt\_Amph | RICH | percent taxa - Order Amphipoda |
| nt\_Isop | RICH | number of taxa - Order Isopoda |
| pi\_Isop | COMP | percent individuals - Order Isopoda |
| pt\_Isop | RICH | percent taxa - Order Isopoda |
| pi\_AmphIsop | COMP | percent individuals - Orders Amphipoda & Isopoda |
| pi\_Baet | COMP | percent individuals - Family Baetidae |
| nt\_Bival | RICH | number of taxa - Class Bivalvia |
| pi\_Bival | COMP | percent individuals - Class Bivalvia |
| pt\_Bival | RICH | percent taxa - Class Bivalvia |
| pi\_Caen | COMP | percent individuals - Family Caenidae |
| nt\_Chiro | RICH | number of taxa - Family Chironomidae |
| pi\_Chiro | COMP | percent individuals - Family Chironomidae |
| pt\_Chiro | RICH | percent taxa - Family Chironomidae |
| nt\_Coleo | RICH | number of taxa - Order Coleoptera |
| pi\_Coleo | COMP | percent individuals - Order Coleoptera |
| pt\_Coleo | RICH | percent taxa - Order Coleoptera |
| nt\_COET | RICH | number of taxa - Order Coleoptera, Odonata, Ephemeroptera & Trichoptera |
| pi\_COET | COMP | percent individuals - Order Coleoptera, Odonata, Ephemeroptera & Trichoptera |
| pt\_COET | RICH | percent taxa - Order Coleoptera, Odonata, Ephemeroptera & Trichoptera |
| nt\_CruMol | RICH | number of taxa - Crustacea & Mollusca |
| pi\_CruMol | COMP | percent individuals - Crustacea & Mollusca |
| nt\_Dipt | RICH | number of taxa - Order Diptera |
| pi\_Dipt | COMP | percent individuals - Order Diptera |
| pt\_Dipt | RICH | percent taxa - Order Diptera |
| nt\_Ephem | RICH | number of taxa - Order Ephemeroptera |
| pi\_Ephem | COMP | percent individuals - Order Ephemeroptera |
| pt\_Ephem | RICH | percent taxa - Order Ephemeroptera |
| pi\_EphemNoCae | COMP | percent individuals - Order Ephemeroptera, excluding Family Caenidae |
| pi\_EphemNoCaeBae | COMP | percent individuals - Order Ephemeroptera, excluding Families Caenidae & Baetidae |
| nt\_EPT | RICH | number of taxa - Orders Ephemeroptera, Plecoptera & Trichoptera (EPT) |
| pi\_EPT | COMP | percent individuals - Orders Ephemeroptera, Plecoptera & Trichoptera (EPT) |
| pt\_EPT | RICH | percent taxa - Orders Ephemeroptera, Plecoptera & Trichoptera (EPT) |
| nt\_Gast | RICH | number of taxa - Class Gastropoda |
| pi\_Gast | COMP | percent individuals - Class Gastropoda |
| pt\_Gast | RICH | percent taxa - Class Gastropoda |
| pi\_Hydro | COMP | percent individuals - Family Hydropsychidae |
| nt\_Insect | RICH | number of taxa - Class Insecta |
| pi\_Insect | COMP | percent individuals - Class Insecta |
| pt\_Insect | RICH | percent taxa - Class Insecta |
| nt\_Mega | RICH | number of taxa - Order Megaloptera |
| pi\_Mega | COMP | percent individuals - Order Megaloptera |
| pt\_Mega | RICH | percent taxa - Order Megaloptera |
| nt\_NonIns | RICH | number of taxa - Class not Insecta |
| pi\_NonIns | COMP | percent individuals - Class not Insecta |
| pt\_NonIns | RICH | percent taxa - Class not Insecta |
| nt\_Odon | RICH | number of taxa - Order Odonata |
| pi\_Odon | COMP | percent individuals - Order Odonata |
| pt\_Odon | RICH | percent taxa - Order Odonata |
| nt\_OET | RICH | number of taxa - Orders Odonata, Ephemeroptera & Trichoptera (OET) |
| pi\_OET | COMP | percent individuals - Orders Odonata, Ephemeroptera & Trichoptera (OET) |
| pt\_OET | RICH | percent taxa - Orders Odonata, Ephemeroptera & Trichoptera (OET) |
| nt\_Oligo | RICH | number of taxa - Class Oligochaeta |
| pi\_Oligo | COMP | percent individuals - Class Oligochaeta |
| pt\_Oligo | RICH | percent taxa - Class Oligochaeta |
| nt\_Pleco | RICH | number of taxa - Order Plecoptera |
| pi\_Pleco | COMP | percent individuals - Order Plecoptera |
| pt\_Pleco | RICH | percent taxa - Order Plecoptera |
| nt\_POET | RICH | number of taxa - Orders Plecoptera, Odonata, Ephemeroptera & Trichoptera (POET) |
| pi\_POET | COMP | percent individuals -Orders Plecoptera, Odonata, Ephemeroptera & Trichoptera (POET) |
| pt\_POET | RICH | percent taxa - Orders Plecoptera, Odonata, Ephemeroptera & Trichoptera (POET) |
| nt\_Trich | RICH | number of taxa - Order Trichoptera |
| pi\_Trich | COMP | percent individuals - Order Trichoptera |
| pt\_Trich | RICH | percent taxa - Order Trichoptera |
| pi\_TricNoHydro | COMP | percent individuals - Order Trichoptera, excluding Family Hydropsychidae |
| pi\_SimBtri | COMP | percent individuals - Families Simuliidae & Baetis tricaudatus |
| pi\_dom01 | RICH | percent individuals - most dominant taxon [max(N\_TAXA)] |
| pi\_dom02 | RICH | percent individuals - two most dominant taxa |
| pi\_dom03 | RICH | percent individuals - three most dominant taxa |
| pi\_dom04 | RICH | percent individuals - four most dominant taxa |
| pi\_dom05 | RICH | percent individuals - five most dominant taxa |
| x\_Shan\_2 | RICH | Shannon Wiener Diversity Index (log base 2) - x\_Shan\_Num/log(2) |
| x\_D | RICH | Simpson's Index |
| x\_Evenness | RICH | Evenness=x\_Shan\_e/log(nt\_total) |
| x\_Becks | TOLER | Becks Biotic Index = 2\*[C1Taxa]+[C2Taxa] (see footnote) |
| x\_HBI | TOLER | Hilsenhoff Biotic Index (references the TolVal field) |
| nt\_tv\_intol | TOLER | number of taxa - tolerance value - intolerant ≤ 3 |
| pi\_tv\_intol | TOLER | percent individuals - tolerance value - intolerant ≤ 3 |
| pt\_tv\_intol | TOLER | percent taxa - tolerance value - intolerant ≤ 3 |
| nt\_tv\_toler | TOLER | number of taxa - tolerance value -tolerant ≥ 7 |
| pi\_tv\_toler | TOLER | percent individuals - tolerance value -tolerant ≥ 7 |
| pt\_tv\_toler | TOLER | percent taxa - tolerance value -tolerant ≥ 7 |
| nt\_ffg\_col | FFG | number of taxa - Functional Feeding Group (FFG) - collector-gatherer (CG) |
| pi\_ffg\_col | FFG | percent individuals - Functional Feeding Group (FFG) - collector-gatherer (CG) |
| pt\_ffg\_col | FFG | percent taxa - Functional Feeding Group (FFG) - collector-gatherer (CG) |
| nt\_ffg\_filt | FFG | number of taxa - Functional Feeding Group (FFG) - collector-filterer (CF) |
| pi\_ffg\_filt | FFG | percent individuals - Functional Feeding Group (FFG) - collector-filterer (CF) |
| pt\_ffg\_filt | FFG | percent taxa - Functional Feeding Group (FFG) - collector-filterer (CF) |
| nt\_ffg\_pred | FFG | number of taxa - Functional Feeding Group (FFG) - predator (PR) |
| pi\_ffg\_pred | FFG | percent individuals - Functional Feeding Group (FFG) - predator (PR) |
| pt\_ffg\_pred | FFG | percent taxa - Functional Feeding Group (FFG) - predator (PR) |
| nt\_ffg\_scrap | FFG | number of taxa - Functional Feeding Group (FFG) - scraper (SC) |
| pi\_ffg\_scrap | FFG | percent individuals - Functional Feeding Group (FFG) - scraper (SC) |
| pt\_ffg\_scrap | FFG | percent taxa - Functional Feeding Group (FFG) - scraper (SC) |
| nt\_ffg\_shred | FFG | number of taxa - Functional Feeding Group (FFG) - shredder (SH) |
| pi\_ffg\_shred | FFG | percent individuals - Functional Feeding Group (FFG) - shredder (SH) |
| pt\_ffg\_shred | FFG | percent taxa - Functional Feeding Group (FFG) - shredder (SH) |
| nt\_habit\_burrow | HABIT | number of taxa - Habit - burrowers (BU) |
| pi\_habit\_burrow | HABIT | percent individuals - Habit - burrowers (BU) |
| pt\_habit\_burrow | HABIT | percent taxa - Habit - burrowers (BU) |
| nt\_habit\_climb | HABIT | number of taxa - Habit - climbers (CB) |
| pi\_habit\_climb | HABIT | percent individuals - Habit - climbers (CB) |
| pt\_habit\_climb | HABIT | percent taxa - Habit - climbers (CB) |
| nt\_habit\_cling | HABIT | number of taxa - Habit - clingers (CN) |
| pi\_habit\_cling | HABIT | percent individuals - Habit - clingers (CN) |
| pt\_habit\_cling | HABIT | percent taxa - Habit - clingers (CN) |
| nt\_habit\_sprawl | HABIT | number of taxa - Habit - sprawlers (SP) |
| pi\_habit\_sprawl | HABIT | percent individuals - Habit - sprawlers (SP) |
| pt\_habit\_sprawl | HABIT | percent taxa - Habit - sprawlers (SP) |
| nt\_habit\_swim | HABIT | number of taxa - Habit - swimmers (SW) |
| pi\_habit\_swim | HABIT | percent individuals - Habit - swimmers (SW) |
| pt\_habit\_swim | HABIT | percent taxa - Habit - swimmers (SW) |
| nt\_volt\_multi | VOLT | number of taxa - multivoltine (MULTI) |
| pi\_volt\_multi | VOLT | percent individuals - multivoltine (MULTI) |
| pt\_volt\_multi | VOLT | percent taxa - multivoltine (MULTI) |
| nt\_volt\_semi | VOLT | number of taxa - semivoltine (SEMI) |
| pi\_volt\_semi | VOLT | percent individuals - semivoltine (SEMI) |
| pt\_volt\_semi | VOLT | percent taxa - semivoltine (SEMI) |
| nt\_volt\_uni | VOLT | number of taxa - univoltine (UNI) |
| pi\_volt\_uni | VOLT | percent individuals - univoltine (UNI) |
| pt\_volt\_uni | VOLT | percent taxa - univoltine (UNI) |
| nt\_ti\_cc | TEMP | number of taxa - thermal indicator - cold/cool |
| pi\_ti\_cc | TEMP | percent individuals - thermal indicator - cold/cool |
| pt\_ti\_cc | TEMP | percent taxa - thermal indicator - cold/cool |
| nt\_ti\_w | TEMP | number of taxa - thermal indicator - warm |
| pi\_ti\_w | TEMP | percent individuals - thermal indicator - warm |
| pt\_ti\_w | TEMP | percent taxa - thermal indicator - warm |

B2 Summary of input metrics in existing low gradient IBIs

Vermont DEC (in progress; personal communication Aaron Moore)

Hybrid low gradient (HLG)

1. Density
2. EOT Richness
3. BCG intolerant richness
4. BCG intolerant COTE %
5. Modified EOT/EOT+Chiro
6. PMA-O
7. Amphipoda+Isopoda %
8. Biotic Index
9. PPCS-F
10. Shr%/CF+Shr%

Soft/slow low gradient (SLG)

1. Density
2. EOT Richness
3. BCG intolerant richness
4. BCG intolerant COTE %
5. Modified EOT/EOT+Chiro
6. PMA-O
7. Amphipoda+Isopoda %
8. Biotic Index
9. PPCS-F
10. Modified EOT Density

New York State DEC (in progress; personal communication Gavin Lemly)

Provisional IBIs by regions for low-gradient streams for three regions:

Great Lakes

rich\_family: decrease with stress

pct\_dom1\_order: increase with stress

shannon\_family: decrease with stress

rich\_scraper: decrease with stress

Adirondacks

pct\_insecta: decrease with stress

rich\_mollusca\_amphipoda\_fa: increase with stress

rich\_intolerant: decrease with stress

rich\_et\_macro\_genspecies: decrease with stress

Hudson Valley+Southern Tier:

pct\_rich\_cote\_family: decrease with stress

pct\_et: decrease with stress

pct\_filterer: decrease with stress

shannon\_genus: decrease with stress

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Location | Index Identifier | Metrics | Metric response to stress | Citation | Remarks |
| Coastal plain region of the 6 states (New Jersey, Delaware, Maryland, Virginia, North Carolina, and South Carolina) | Coastal Plain Macroinvertebrate Index (CPMI) | # of taxa  # of EPT taxa  % Ephemeroptera  Hilsenhoff Biotic Index (HBI)  % clingers | **# taxa**: decrease; 45% overallassessment accuracy  **# EPT taxa**: decrease; high assessment accuracy (84% overall); correlated with Ephem and Trichop metrics; historic reliability.  **% Ephemeroptera**: decrease; 57% overallassessment accuracy; lower redundancies with HBI & # EPT; high redundancy with % EPT metrics already selected  **HBI**: increase; high assessment accuracy (80%overall); strongly correlated with other tolerance metrics; historic reliability.  **% clingers**: decrease; not redundant with TT and %E metrics already selected, moderately redundant with the HBI and EPT metrics | Maxted et al. 2000 | Oct/Nov sampling period.  Accurately identified 86% of impaired sites overall (varied 83-100% across the 3 regions classified).  90% CI for the 5 core metrics were +6.0 taxa for TT, +2.5 taxa for EPT, +8.9% for %E, +0.28 units for the HBI, +13.8% for %CL, and +3.1 units for the CPMI. |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Location | Index Identifier | Metrics | Metric response to stress | Citation | Remarks |
| Florida | Stream Condition  Index (SCI) | # total taxa  # EPT taxa  # Chironomidae taxa  Florida Index  % dominant taxa  % Diptera  % gatherers  % filterers | **# total taxa**: decrease  **# EPT taxa**: decrease  **# Chironomidae taxa**: decrease  **Florida Index**: decrease  **% dominant taxa**: increase  **% Diptera**: increase  **% gatherers**: variable  **% filtere**rs: decrease; “filter feeders are also thought to be sensitive in low-gradient streams (Wallace et al. 1977).” | Barbour et al. 1996 | Summer index sampling period (Jul-Sep).  3 classified regions: panhandle, peninsular Florida, & the northeastern portion of Florida  Scores (5, 3, or 1) developed for 8 metrics to allow aggregation into an index |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Location | Index Identifier | Metrics | Metric response to stress | Citation | Remarks |
| Choctawhatchee-Pea Rivers watershed, AL | Invertebrate community index (ICI) | # EPT taxa  # Trichoptera taxa  # Diptera taxa  # Crustacea + Mollusca  % Dominant taxa  % Ephemeroptera  % Diptera  % Chironominae to chironomids  Family Biotic Index (FBI)  % Shredders | # EPT taxa: decrease  # Trichoptera taxa: decrease  # Diptera taxa: decrease  # Crustacea + Mollusca: decrease  % Dominant taxa: increase  % Ephemeroptera: decrease  % Diptera: increase  % Chironominae to chironomids: decrease  Family Biotic Index (FBI): increase  % Shredders: decrease | Bennet et al. 2004 | Within the coastal plains ecoregion in southeast Alabama; low elevation and loosely compacted, sandy soils.  34 wadeable first through sixth-order streams; plus for validation 7 additional least impacted and 8 impacted streams.  49 sites sampled once during April and May 2001.  The 10 selected metrics (of 38 tested) had significant correlations with one or more physiochemical variables.  ICI calculated by summing the 10 metric scores from 34 sites; ranged from 18 to 56 out of a possible score of 60.  The ICI was not always capable of discriminating between artificially enriched sites and good quality sites |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Location | Index Identifier | Metrics | Metric response to stress | Citation | Remarks |
| Central Valley, CA | Central Valley IBI | collector richness  predator richness  percent EPT taxa  percent clinger taxa  Shannon diversity | collector richness: decrease  predator richness: decrease  percent EPT taxa: decrease  percent clinger taxa: decrease  Shannon diversity: decrease  *Note: these expectations deduced from the scoring ranges presented in Table 2 of paper.* | Rehn et al. 2008 | Perennial streams on the valley floor  In the Central Valley, minimally disturbed reference sites no longer available.  Most streams are highly alteredby human activities such as urbanization, agriculture and water diversions.  80 metrics evaluated; metric criteria: 1) sufficient range for scoring; 2) responsiveness to land use and reach-scale disturbance variables (as data allowed); 3) good discrimination between reference and test sites; 4) lack of correlation with other responsive metrics.  Lack of intolerant and shredder taxa in Valley floor streams.  Final IBI more strongly related to reach-scale physical habitat variables than to water chemistry or land use variables.  The final 5 IBI metrics did not vary between spring and fall samples and did not require seasonal adjustments in scoring. |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Location | Index Identifier | Metrics | Metric response to stress | Citation | Remarks |
| Austria | Multimetric index (for A01 - Mid-sized (low-gradient) streams in the Hungarian Plains) | # of total families  # of EP # taxa  # of Plecoptera (abundance)  [%] EP # individuals  [%] EP # taxa  Saprobic index  # of sensitive taxa  [%] Shredder  Diversity (Margalef) | # of total families: decrease  # of EP # taxa: decrease  # of Plecoptera (abundance):  [%] EP # individuals: decrease  [%] EP # taxa: decrease  Saprobic index: increase  # of sensitive taxa: decrease  [%] Shredder: decrease  Diversity (Margalef): decrease | Ofenböck et al. 2004 | Stressor – organic pollution |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Location | Index Identifier | Metrics | Metric response to stress | Citation | Remarks |
| Willow Creek, Nebraska | Composite Biotic Index (CBI) | Percent dominance  EPT index (i.e. # EPT taxa)  EPT abund/EPT + chironomid abundance  Scraper abund/filterer abund  Taxa richness  Hilsenhoff index | Not specified in paper | Whiles et al. 2002 | Developed with metrics used previously by the NDEQ during their statewide stream survey (NDEQ 1991).  CBI scores actually are based on a “reference condition” for Nebraska rather than the reference stream (site 4) in our basin  Corrected metrics for stream size (based on discharge) using relationships generated from a prior investigation; i.e. metrics were scored 1, 3, or 5 based on regression equations generated by the NDEQ (1991) that divided scatter plots of stream size vs metric scores into thirds |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Location | Index Identifier | Metrics | Metric response to stress | Citation | Remarks |
| Netherlands | Multimetric index (for slow-running streams) | See metrics listed in Table 2 copied below from paper. | complex | Vlek et al. 2004 | Included metrics that indicated the different classes (from 5 (high quality) to 1 (low quality); final index equation combined these; for slow running streams:    Where:  S, final score;  T1, sum of scores for the individual metrics indicating class 1; T2, sum of scores for the individual metrics indicating class 2; etc. And n1, number of indices indicating class 1; etc.  Validation showed that 54% of the streams were classified correctly |

References – existing low gradient IBIs

Barbour, M.T., J. Gerritsen, G.E. Griffith, R. Frydenborg, E. McCarron, J.S. White, and M.L. Bastoam. 1996. A framework for biological criteria for Florida streams using benthic macroinvertebrates. Journal of the North American Benthological Society 15:185–211.

Bennett, H.H., M.W. Mullen, P.M. Stewart, J.A. Sawyer, and E.C. Webber. 2004. Development of an invertebrate community index for an Alabama coastal plain watershed. Journal of the American Water Resources Association 40(1): 43-51. https://doi.org/10.1111/j.1752-1688.2004.tb01008.x|.

Lemley, G.M., B.T. Duffy, Z.M. Smith, A.J. Smith, and R.G. Smith. 2019. Assessing Low Gradient Streams in New York State: Developing a Macroinvertebrate IBI. PowerPoint Presentation, February 28th, 2019. New York Department of Environmental Conservation.

Maxted, J.R., M.T. Barbour, J. Gerritsen, V. Poretti, N. Primrose, A. Silvia, D. Penrose, and R. Renfrow. 2000. Assessment framework for mid-Atlantic coastal plain streams using benthic macroinvertebrates. Journal of the North American Benthological Society 19(1):128–144.

Mazor, R.D., K.C. Schiff, K.J. Ritter, A. Rehn, and P. Ode. 2009. Bioassessment tools in novel habitats: An evaluation of indices and sampling methods in low-gradient streams in California. Environmental Monitoring and Assessment 167:91–104.

Ofenböck. T., O. Moog, J. Gerritsen, and M. Barbour. 2004. A stressor specific multimetric approach for monitoring running waters in Austria using benthic macro-invertebrates. Hydrobiologia 516:251–268.

Rehn, A.C., J.T. May, and P.R. Ode. 2008. An Index of Biotic Integrity (IBI) for Perennial Streams in California’s Central Valley. Surface Water Ambient Monitoring Program, California Water Boards, Technical Report, 33 pp.

Vlek, H.E., Piet F. M. Verdonschot & Rebi C. Nijboer. 2004. Towards a multimetric index for the assessment of Dutch streams using benthic macroinvertebrates. Hydrobiologia 516:173–189.

Whiles, M.R., B.L. Brock, A.C. Franzen, and S.C. Dinsmore. 2004. Stream Invertebrate Communities, Water Quality, and Land-Use Patterns in an Agricultural Drainage Basin of Northeastern Nebraska, USA. Environmental Management 26(5):563–576. DOI: 10.1007/s002670010113.

B3 Non-target taxa

The following non-target taxa were excluded from metric calculations:

|  |  |  |
| --- | --- | --- |
| ORDER | TAXAID | NONTARGET |
| Hemiptera | Belostoma | TRUE |
| Hemiptera | Belostomatidae | TRUE |
| Hemiptera | Corixidae | TRUE |
| Hemiptera | Gerridae | TRUE |
| Hemiptera | Gerris | TRUE |
| Hemiptera | Microvelia | TRUE |
| Hemiptera | Neoplea striola | TRUE |
| Hemiptera | Notonecta | TRUE |
| Sarcoptiformes | Oribatida | TRUE |
| Hemiptera | Pleidae | TRUE |
| Hemiptera | Ranatra | TRUE |
| Hemiptera | Rhagovelia | TRUE |
| Hemiptera | Veliidae | TRUE |

For the purposes of IBI calculations, “macroinvertebrate” is defined to include:

         all aquatic Annelida;

         all aquatic Mollusca;

         aquatic macro Crustacea (except as noted below);

         all aquatic Arachnida except for Oribatid mites (which are not truly aquatic); and

         the aquatic life stages of Insecta except Hemiptera and adult Coleoptera other than Elmidae.

Those macroinvertebrates excluded from the above list are not used for one of three reasons: either there is insufficient ecological information on them to make them useful for biomonitoring, they are surface film dwellers, or they are capable of escaping the aquatic environment at will to avoid temporarily unfavorable conditions.  One further exception is crayfish (Class Crustacea, Family Cambaridae), which often are seen evacuating the immediate area as kick-sampling begins, and even swimming out of the kick-net.  Crayfish species are noted when present in the sample but are not counted toward total numbers.

B4 Exclusion criteria for redundant taxa

When calculating metrics for benthic macroinvertebrates, there are occasions when certain taxa are not included in taxa richness metrics but the individuals are included for all other metrics. This is done to avoid double counting taxa that may have been identified to a more coarse level when taxa of a finer level are present in the same sample.

These taxa have been referred to by many names – e.g., Excluded Taxa, NonUnique Taxa, or Ambiguous Taxa. This document will use the term Excluded.

We used the ‘markExcluded’ function in the BioMonTools R package (<https://github.com/leppott/BioMonTools>) to mark redundant taxa in the low gradient samples prior to metric calculations. Redundant taxa were identified on a sample-by-sample basis and excluded from the richness calculations.

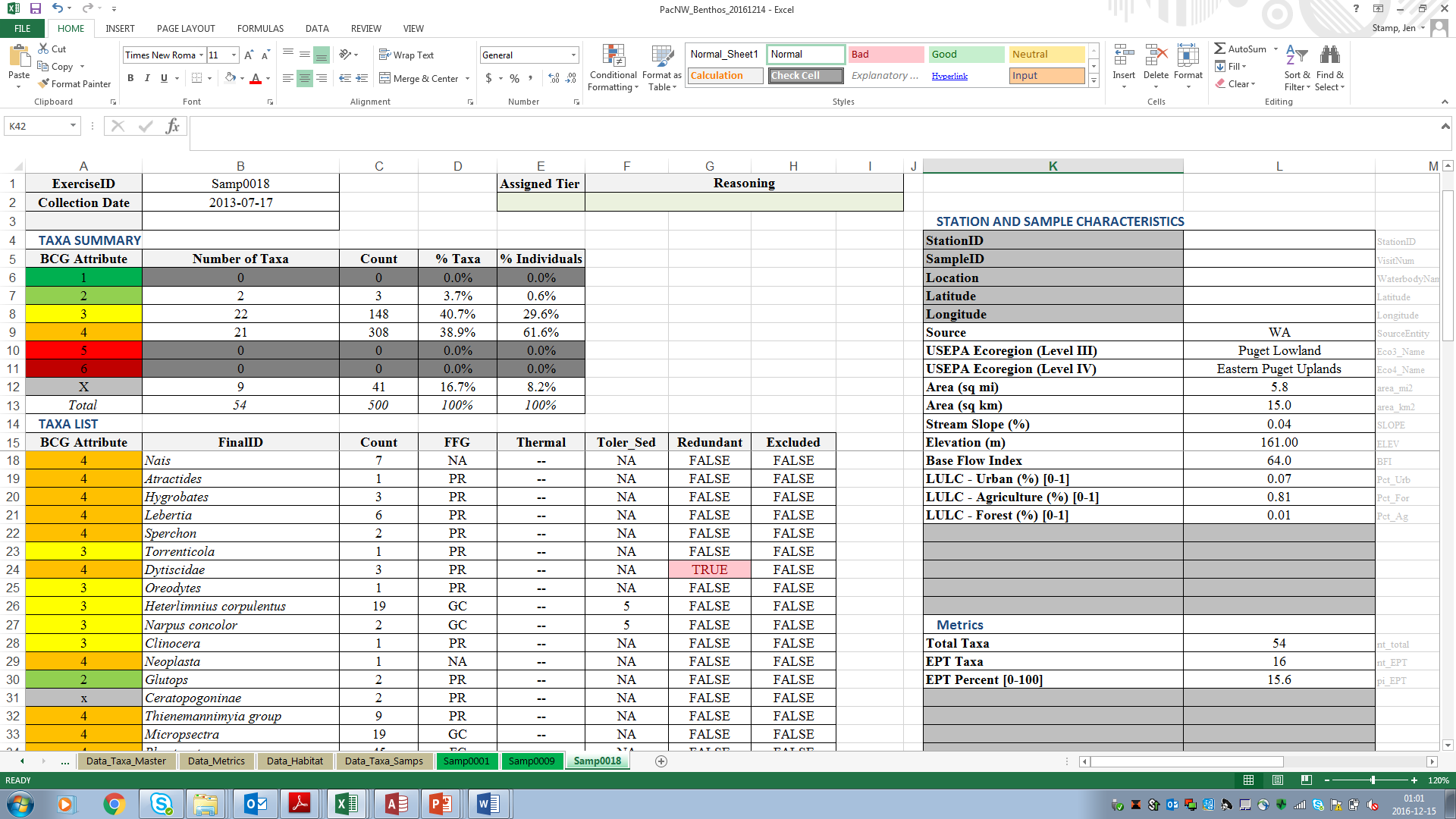
Redundant taxa were identified based on the following steps:

1. Calculate and find all taxa names that appear in a sample at each taxonomic rank more than once (for an example, see Figure 1). These are the potential "parents" to be excluded.

2. Check if any of the potential "parents" equal a final ID in their respective samples.

3. If you get a match these are marked as "Excluded"

All Excluded decisions are sample-specific and the rules should be reapplied if sample contents change. Also, if the level of effort or operational taxonomic units change, the Excluded taxa designations should be recalculated.



**Figure 1**. Example - Dytiscidae (family-level) is excluded from the richness metrics in this sample because these organisms could be the same taxon as Oreodytes (genus-level). The exclusion rule is applied on a sample by sample basis.

Below is a more detailed description of the process that the markExcluded function follows. Before starting, it is necessary to have a complete and correct master taxa list (all phylogenetic information and ranks).

## Terminology

* Target Rank = intended level of taxonomy for identification, e.g., genus. Typically, specified in the project’s SOP but can be adjusted during the OTU process.
* Parent or Parent Taxon = a taxon that occurs in the data in addition to other taxa in the same group that are identified to a more specific level. For example, the family Baetidae may occur in the data in addition to genera within the family Baetidae. In this case the name Baetidae is a parent to the other taxa within the family. Parents do not have to be only a single rank above the child taxon. That is, the class and order ranks are parents of any family ranks within them.
* Child or Children Taxa = a taxa or taxon that occurs in the data in addition to individuals identified to a coarser level. For example, the genera Baetis and Procloeon may occur in addition to the family Baetidae (of which the 2 genera listed are a member). In this case Baetis and Procloen are children of Baetidae.

## Rule Development

For each sample:

1. Determine “potential” taxa for exclusion based on rank (or level) names appearing more than once in a sample.
   1. This is done for all ranks present; phylum, class, order, family, tribe, genus, species.
2. Check if any “potential” taxa are equal to a final (unique) ID in the same sample.
3. Stage is combined with taxa names if used in the dataset.

## Requirements

1. A sample taxa table or data frame.
   1. All non-count and zero individual taxa have been removed.
   2. Unique sample ID code in a single column.
   3. A column with a final identification that is narrative not numeric. That is, Baetidae is ok but the ITIS number is not.
   4. Phylogenetic rank/level columns.
      1. This can be applied from a master taxa table but needs to be included in this table. One column per rank.
      2. Names need to be consistently spelled.

## Procedures

1. Find all potential Parents (those with a rank coarser than the target rank). This is done by creating a list of taxa rank names that appear more than once in a sample. This is done for each taxonomic rank.
2. The above list is compared to the final identifications for each sample.
   1. Special consideration is made for ranks of finer detail than genus. That is, names that are a combination of more than one field.
3. Any matches are marked as “Excluded”.

There is still a need for manual review / QC check of the final list of Excluded designations.