NEON USER GUIDE TO AQUATIC PLANT, BRYOPHYTE, LICHEN, AND MACROALGAE POINT COUNTS IN WADEABLE STREAMS (NEON.DP1.20072)

PREPARED BY	ORGANIZATION	DATE
Stephanie Parker	AQU	07/28/2017
Caren Scott	AQU	07/28/2017

CHANGE RECORD

REVISION	DATE	DESCRIPTION OF CHANGE
A	07/03/2017	Initial Release

TABLE OF CONTENTS

LIST OF TABLES AND FIGURES

1 DESCRIPTION

1.1 Purpose

This document provides an overview of the data included in this NEON Level 1 data product, the quality controlled product generated from raw Level 0 data, and associated metadata. In the NEON data products framework, the raw data collected in the field, for example, the point count field samples from a single collection event are considered the lowest level (Level 0). Raw data that have been quality checked via the steps detailed herein, as well as simple metrics that emerge from the raw data are considered Level 1 data products.

The text herein provides a discussion of measurement theory and implementation, data product provenance, quality assurance and control methods used, and approximations and/or assumptions made during L1 data creation.

1.2 Scope

This document describes the steps needed to generate the L1 data product, Aquatic plant, bryophyte, lichen, and macroalgae point counts in wadeable streams and associated metadata from input data. This document also provides details relevant to the publication of the data products via the NEON data portal, with additional detail available in the file, NEON Data Variables for Aquatic plant, bryophyte, lichen, and macroalgae point counts in wadeable streams (NEON.DP1.20072) (AD[04]), provided in the download package for this data product.

This document describes the process for ingesting and performing automated quality assurance and control procedures on the data collected in the field pertaining to AOS Protocol and Procedure: Aquatic plant, bryophyte, lichen, and macroalgae sampling (AD[06]). The raw data that are processed in this document are detailed in the file, NEON Raw Data Validation for Aquatic plant, bryophyte, lichen, and macroalgae point counts in wadeable streams (NEON.DP0.20072) (AD[03]), provided in the download package for this data product. Please note that raw data products (denoted by 'DP0') may not always have the same numbers (e.g., '10033') as the corresponding L1 data product.

2 RELATED DOCUMENTS AND ACRONYMS

2.1 Associated Documents

AD[01]	NEON.DOC.000001	NEON Observatory Design (NOD) Requirements
AD[02]	NEON.DOC.002652	NEON Level 1, Level 2 and Level 3 Data Products Catalog
AD[03]	NEON.DP0.20072.001 _dataValidation.csv	NEON Raw Data Validation for Aquatic plant, bryophyte, lichen, and macroalgae point counts in wadeable streams (NEON.DP0.20072)
AD[04]	NEON.DP1.20072.001 _variables.csv	NEON Data Variables for Aquatic plant, bryophyte, lichen, and macroalgae point counts in wadeable streams (NEON.DP1.20072)
AD[05]	NEON.DOC.001152	NEON Aquatic Sampling Strategy
AD[06]	NEON.DOC.003039	AOS Protocol and Procedure: Aquatic plant, bryophyte, lichen, and macroalgae sampling
AD[07]	NEON.DOC.000008	NEON Acronym List
AD[08]	NEON.DOC.000243	NEON Glossary of Terms
AD[09]	OS_Generic_Transitions .pdf	NEON Algorithm Theoretical Basis Document: OS Generic Transitions
AD[10]	Nicl Language.pdf	NEON's Ingest Conversion Language (NICL) specifications

2.2 Acronyms

Acronym	Definition
USDA	US Department of Agriculture
USEPA	US Environmental Protection Agency
USGS	US Geological Survey

3 DATA PRODUCT DESCRIPTION

Point counts of aquatic plants, bryophytes, lichens, and macroalgae include taxonomic identification and relative abundance of plants. These data address the NEON Grand Challenge areas of Biodiversity, Biogeochemistry, and Invasive Species, as well provide important data on the macrophyte community composition in wadeable streams, which may be used to assess the status of the aquatic ecosystem. Aquatic vegetation is sampled three times per year at each NEON aquatic site (AD[05]). Sampling dates are based on a combination of variables, including hydrology in streams, accumulated degree days (temperature), and riparian greenness (phenology). For additional information see the sampling design NEON Aquatic Sampling Strategy (AD[05]) and protocol AOS Protocol and Procedure: Aquatic plant, bryophyte, lichen, and macroalgae sampling (AD[06]).

3.1 Spatial Sampling Design

Point count data are collected at 10 permanently-marked locations per wadeable stream site. Each transect (Figure 1) is established during initial sampling at the site. Five transects are established in the dominant habitat type, and five in the second-most dominant habitat type. Transects should be established in alternating habitat types along the reach (e.g., transect 1 in a pool, transect 2 in a riffle, transect 3 in a pool). If it is not possible to separate each transect by a different habitat type, transects may be spaced at least 10 m apart within the same habitat type/unit. Transects are established during the first sampling bout by domain technicians and coordinates are collected using a high-precision GPS unit (e.g., Trimble with a tornado antenna). Plot markers are established on one bank, with the transect extending perpendicular to stream flow across the wetted channel. Transects remain in place until significant morphological changes occur in the stream (e.g., bed-moving spate) that causes the transect to be unusable. Only if the geomorphology of the site changes significantly will a transect re-established. See AOS Protocol and Procedure: Aquatic plant, bryophyte, lichen, and macroalgae sampling (AD[06]) for additional details on strategy and SOPs.

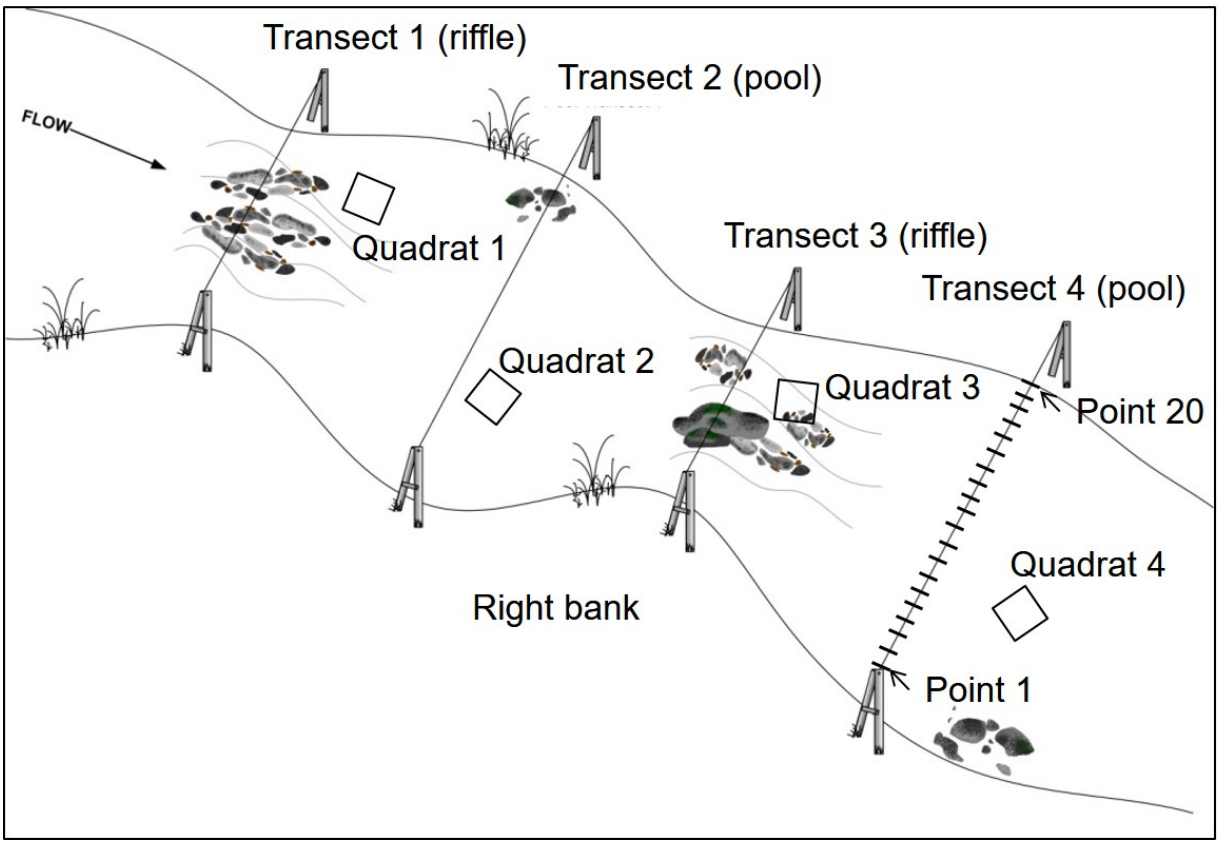

Figure 1: Aquatic plant sampling locations in wadeable streams, showing quadrat locations in proximity to established transects

3.2 Temporal Sampling Design

Aquatic point count data are collected during all three biological and sediment chemistry sampling bouts (roughly spring, summer, and fall). Data collection for this data product occurs within one day during a sampling bout at a given site. The number of specimens produced depends on the density of vegetation at the site.

3.3 Variables Reported

All variables reported from the field or laboratory technician (L0 data) are listed in the file, NEON Raw Data Validation for Aquatic plant, bryophyte, lichen, and macroalgae point counts in wadeable streams (NEON.DP0.20072) (AD[03]). All variables reported in the published data (L1 data) are also provided separately in the file, NEON Data Variables for Aquatic plant, bryophyte, lichen, and macroalgae point counts in wadeable streams (NEON.DP1.20072) (AD[04]).

Field names have been standardized with Darwin Core terms (http://rs.tdwg.org/dwc/; accessed 16 February 2014), the Global Biodiversity Information Facility vocabularies (http://rs.gbif.org/vocabulary/gbif/; accessed 16

February 2014), the VegCore data dictionary (https://projects.nceas.ucsb.edu/nceas/projects/bien/wiki/VegCore; accessed 16 February 2014), where applicable. NEON Terrestrial Observation System (TOS) spatial data employs the World Geodetic System 1984 (WGS84) for its fundamental reference datum and Earth Gravitational Model 96 (EGM96) for its reference gravitational ellipsoid. Latitudes and longitudes are denoted in decimal notation to six decimal places, with longitudes indicated as negative west of the Greenwich meridian.

Some variables described in this document may be for NEON internal use only and will not appear in downloaded data.

3.4 Temporal Resolution and Extent

The finest resolution at which temporal data are reported is at **collectDate**, the date and time of day when the sample was collected in the field.

The finest temporal resolution that Aquatic Plant Bryophyte Macroalgae Point Count data will be tracked is at the level of an individual observation per sampling bout. The total number of bouts per year is 3 per aquatic site. Each sampling bout occurs within one day at a given site. The number of observations produced depends on the density of vegetation at the site. Data collected for Aquatic plant bryophyte macroalgae clip harvest (NEON.DOM.SITE.DP0.20066) are collected at the same time.

3.5 Spatial Resolution and Extent

The finest resolution at which spatial data are reported is near a single transect in wadeable streams).

locationID (ID of transect within site) -> siteID (ID of NEON site) -> domainID (ID of a NEON domain)

The basic spatial data included in the data downloaded include the latitude, longitude, and elevation of the named location at the aquatic site (e.g., the permanent transect location in a wadeable stream).

3.6 Associated Data Streams

Data for the Aquatic plant bryophyte macroalgae clip harvest product (NEON.DP1.20066) are collected alongside point count data during bout 2. The two data products may be linked through their shared **eventID** (or a combination of **siteID** and **collectDate**). In addition, morphospecies names and identifications for all aquatic plants products (including this one) are provided in the apc_morphospecies table, provided in the expanded package with downloads of either product and can be linked through **siteID**, **morphospeciesCreatedDate**, and **morphospeciesID**.

Data and maps produced by the Bathymetric and morphological maps data product (NEON.DP4.00132) may also be relevant to point count data.

3.7 Product Instances

There are a maximum of 3 point count bouts per year per site, with counts occurring at each of 10 transects. The number of points observed during each sampling event depends on the width of the stream.

3.8 Data Relationships

A record in apc_perTaxon, apc_taxonomyRaw, or apc_taxonomyProcessed must have a corresponding record in apc_pointTransect. Data in apc_pointTransect are field data that apply to the transect as a whole, A record in apc_pointTransect exists for each point visited along a stream transect, whether a plant taxon was present or not; apc_perTaxon records are only created for points where plants are found. apc_taxonomy records exist only where physical samples have been collected as part of apc_perTaxon sampling (non-null **sampleID**). Duplicates and/or missing data may exist where protocol and/or data entry aberrations have occurred; users should check data carefully for anomalies before joining tables.

apc_pointTransect.csv - > Data are recorded for all points surveyed along the wadeable stream transect in this table.

apc_perTaxon.csv - > If **targetTaxaPresent** = "Yes" in apc_pointTransect, a record will be entered here. All vegetation recorded at the point are recorded here, typically that is for one taxon, but may be for multiple taxa if there is vertical stacking of plants at the point. Data are linked to apl_pointTransect through **namedLocation** and **pointID**.

apc_taxonomyRaw - > One record is created for each **sampleID** in apc_perTaxon that is identified to the lowest practical taxonomist level. The raw taxonomy table represents taxonomy from the domain support facility or the external lab, and is not checked against the NEON taxonomy table for plants or algae. Data are linked to apc_perTaxon through **sampleID**.

apc_taxonomyProcessed.csv - > One record is created for each physically collected **sampleID** in apc_perTaxon that is identified to the lowest practical taxonomist level. The processed taxonomy table represents taxonomy from the domain support facility or the external lab that has been checked and standardized against the NEON taxonomy table for plants or algae. Data are linked to apc_perTaxon through **sampleID**.

apc_voucher.csv - > Plant voucher data may be collected at any time during the field season, when field technicians see an appropriate specimen in the field. Data apply to the NEON aquatic site, and may be linked through **taxonID** and **siteID**. A **sampleID** is created during voucher field collection that can be linked to the external taxonomy data (apc_voucherTaxonomyRaw and apc_voucherTaxonomyProcessed).

apc_voucherTaxonomyRaw - > One record is created for each physically collected **sampleID** that is identified to the lowest practical taxonomist level. The raw taxonomy table represents taxonomy from the domain support facility or the external lab, and is not checked against the NEON taxonomy table for plants or algae. Data are linked to apc_voucher through **sampleID**.

apc_voucherTaxonomyProcessed.csv - > One record is created for each physically collected **sampleID** that is identified to the lowest practical taxonomist level. The processed taxonomy table represents taxonomy from the domain support facility or the external lab that has been checked and standardized against the NEON taxonomy table for plants or algae. Data are linked to apc_voucher through **sampleID**.

apc_morphospecies.csv - > The morphospecies table is used both for this data product, and the clip harvest data product (NEON.DP1.20066). Morphospecies are assigned when the specimen cannot be identified by the NEON field technician. Technicians may revisit this table to update **taxonID** data at a later time if they are able to positively identify the specimen. New morphospecies names are assigned per site each year. Not all morphospecies may be resolved.

3.9 Special Considerations

Percent cover may be calculated using the equation from Bowden et al. 2006:

$$percentCover_i = \frac{N_i}{N_t} \quad 100 \tag{1}$$

Where N_i is the number of observed points in a transect that match class type "i" (i.e., a particular **taxonID**) and N_t is the total number of points observed in the transect. This calculation can generate percent cover values >100% if there is vertical stacking of plants.

4 DATA ENTRY CONSTRAINT AND VALIDATION

Many quality control measures are implemented at the point of data entry within a mobile data entry application or web user interface (UI). For example, data formats are constrained and data values controlled through the provision of dropdown options, which reduces the number of processing steps necessary to prepare the raw data for publication. The data entry workflow for collecting aquatic plant clip harvest field data is diagrammed in Figure 2, and the lab workflow is diagrammed in Figure 3. An additional set of constraints are implemented during the process of ingest into the NEON database. The product-specific data constraint and validation requirements built into data entry applications and database ingest are described in the document NEON Raw Data Validation for Aquatic plant, bryophyte, lichen, and macroalgae point counts in wadeable streams (NEON.DP0.20072), provided with every download of this data product. Contained within this file is a field named 'entryValidationRulesForm', which describes syntactically the validation rules for each field built into the data entry application. Data entry constraints are described in Nicl syntax in the validation file provided with every data download, and the Nicl language is described in NEON's Ingest Conversion Language (NICL) specifications ([AD[10]).

5 DATA PROCESSING STEPS

Following data entry into a mobile application of web user interface, the steps used to process the data through to publication on the NEON Data Portal are detailed in the NEON Algorithm Theoretical Basis Document: OS Generic Transitions (AD[10]).

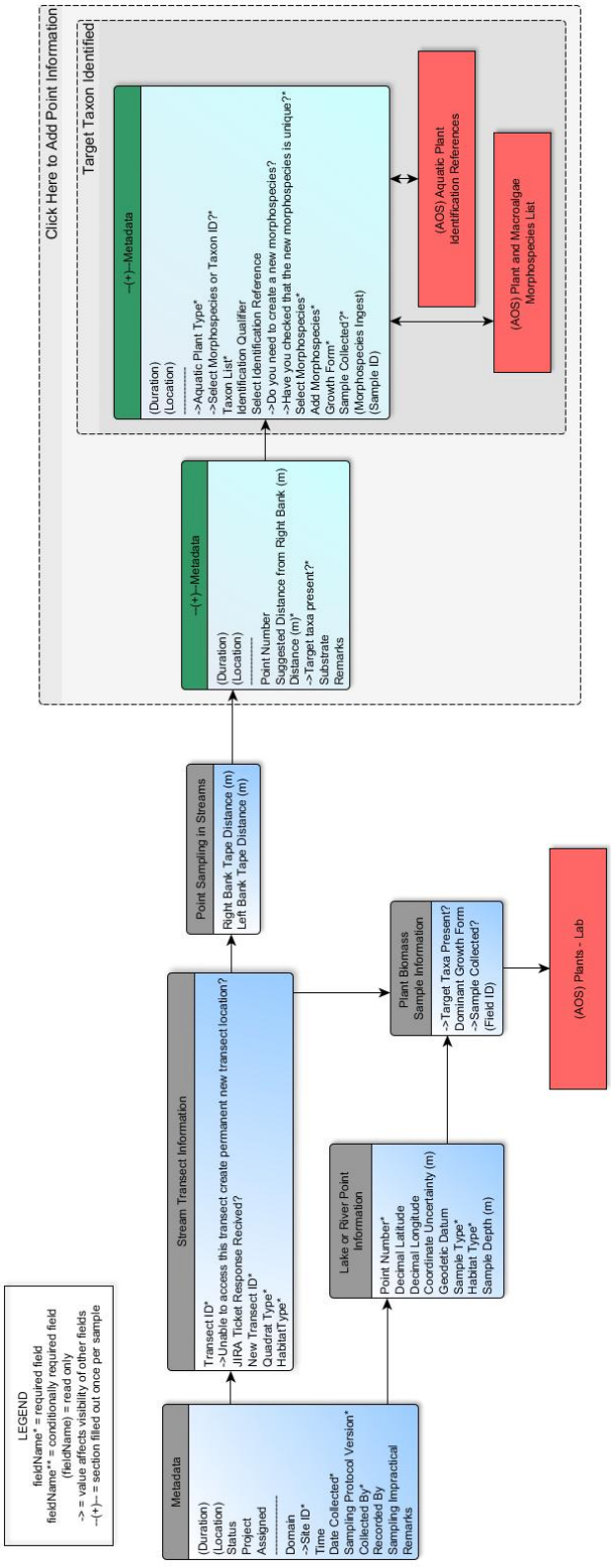

Figure 2: Schematic of the applications used by field technicians to enter aquatic point count and clip harvest field data

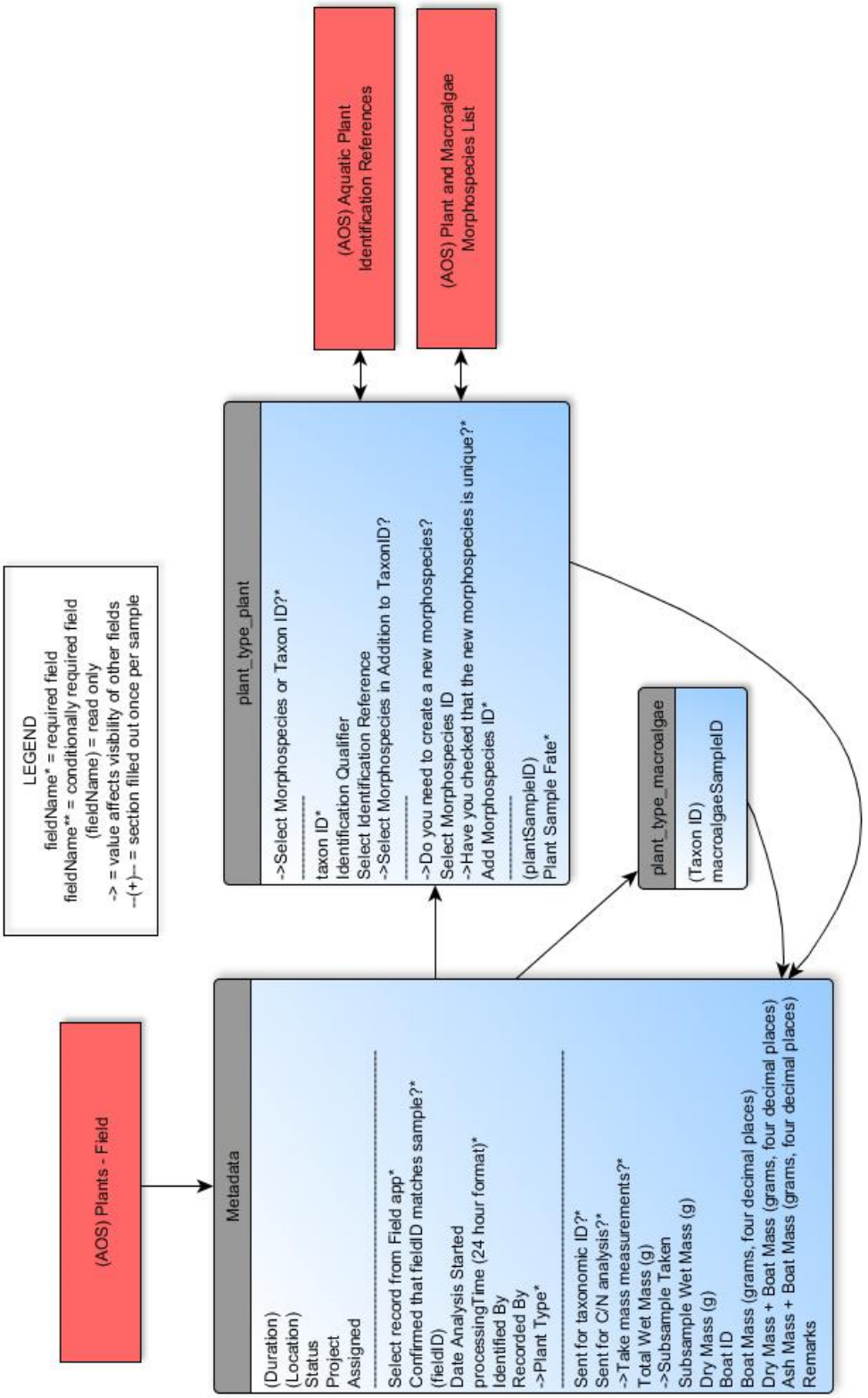

Figure 3: Schematic of the applications used by field technicians to measure aquatic plant biomass data in the domain lab

6 REFERENCES

Bowden, W.B., J.M. Glime, and T. Riis. 2006. Macrophytes and Bryophytes. Pages 381-414 in F.R. Hauer and G.A. Lamberti, editors. Methods in Stream Ecology, Second Edition. Elsevier, Inc., Boston, Massachusetts, USA.