Background on Fuzzy Set Theory

BCG level descriptions in the conceptual model are intentionally general (e.g., "reduced richness") to reflect shared patterns of biological change to human disturbance across waterbody types and ecological regions. To allow for consistent assignments of sites to levels, it is necessary to formalize the expert knowledge by codifying level descriptions into a set of rules that replicate the decision criteria of the expert panel (e.g., Droesen 1996). People tend to use strength of evidence in defining decision rules, and in allowing some deviation from their ideal for any individual attributes, as long as most attributes are in or near the desired range. For example, the definitions of "high," "moderate," "low," etc., are qualitative (but ordinal) and can be interpreted and measured to mean different things. An important step in the BCG process is development of expert consensus defining these, or other, general terms and documenting the expert logic that is the basis for the decisions. The decision rules preserve the collective professional judgment of the expert group and set the stage for the development of models that can reliably assign sites to levels without having to reconvene the same group. In essence, the rules and the models capture the panel's collective decision criteria.

Quantification of rules allows users to consistently assess sites according to the same rules used by the expert panel, and allows a computer algorithm, or other persons, to obtain the same level assignments as the panel. BCG quantitative models have been constructed for over ten different regions based on modern mathematical set theory and logic (called "fuzzy set theory"). Fuzzy set theory is directly applicable to environmental assessment and has been used extensively in engineering applications worldwide (e.g., Demicco and Klir 2004) and environmental applications have been explored in Europe and Asia (e.g., Castella and Speight 1996, Ibelings et al. 2003).

Mathematical fuzzy set theory allows degrees of membership in sets, and degrees of truth in logic, compared to all-or-nothing in classical set theory and logic. Membership of an object in a set is defined by its membership function, a function that varies between 0 and 1. To illustrate, we compare how classical set theory and fuzzy set theory treat the common classification of sediment, where sand is defined as particles less than or equal to 2.0 mm diameter, and gravel is greater than 2.0 mm (Demicco and Klir 2004). In classical "crisp" set theory, a particle with diameter of 1.999 mm is classified as "sand", and one with 2.001 mm diameter is classified as "gravel." In fuzzy set theory, both particles have nearly equal membership (approximately 0.5) in both classes (Demicco and Klir 2004). Very small measurement error in particle diameter greatly increases the uncertainty of classification in classical set theory, but not in fuzzy set theory (Demicco and Klir 2004). Demicco and Klir (2004) proposed four reasons why fuzzy sets and fuzzy logic enhance scientific methodology:

- Fuzzy set theory has capability to deal with "irreducible measurement uncertainty," as in the sand/gravel example above.
- Fuzzy set theory captures vagueness of linguistic terms, such as "many," "large" or "few."
- Fuzzy set theory and logic can be used to manage complexity and computational costs of control and decision systems.

- Fuzzy set theory attempts to model human reasoning and decision-making, which is critically important for defining thresholds and decision levels for environmental management.

The BCG models use mathematical fuzzy logic to replicate human reasoning. Each linguistic variable (e.g., "high taxa richness") is defined quantitatively as a fuzzy set (e.g., Klir 2004). Lower and upper ("fuzzy set") bounds are set for each metric based on distributions of biological metrics across BCG levels. Each metric receives a membership value ranging from 0 to 1, depending on where the value falls in relation to the bounds. The rule threshold falls in the middle of these bounds. Metric values that are less than or equal to the lower bound receive a membership value of 0, while metric values that are greater than or equal to the upper bound receive a membership value of 1. In the example shown below (Figure 1), the example rule for total taxa richness is ≥ 20 (15-25) (the lower bound is 15 and the upper bound is 25), which means –

- If there are 15 or fewer total taxa in the sample, the metric membership value is 0.
- If there are 25 or more total taxa in the sample, the metric membership value is 1.
- If the number of total taxa falls within the lower and upper bounds, the metric membership value will range from 0 to 1 (e.g., if there are 20 total taxa, the membership value will be 0.5; if there are 17 total taxa, the membership value will be 0.2; if there are 23 total taxa, the membership value will be 0.8).

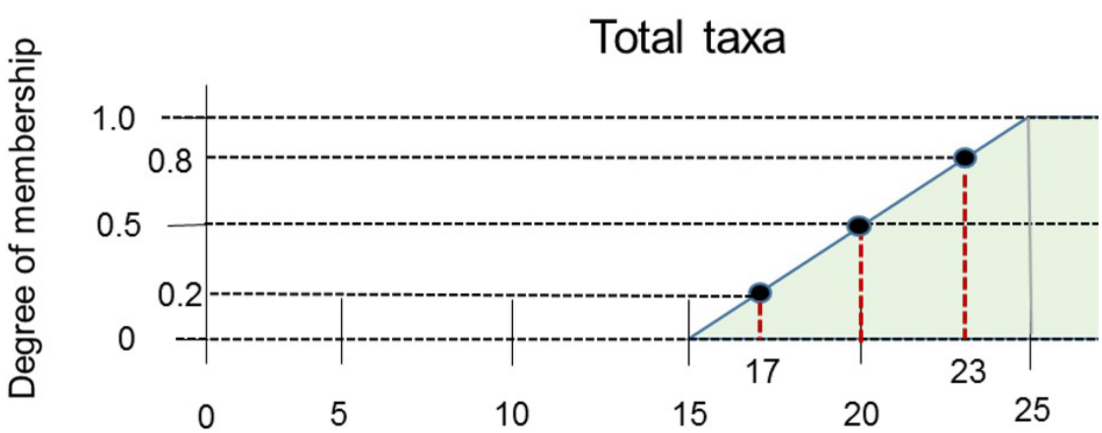

Figure 1. Illustration of the lower and upper ("fuzzy set") bounds for an example metric, total taxa richness. Each metric receives a membership value ranging from 0 to 1, depending on where the value falls in relation to the bounds. In this example, the BCG rule for total taxa richness is ≥ 20 (15-25) (the lower bound is 15 and the upper bound is 25). The black dots show examples of metric membership values assigned to different metric values (e.g., if there are 20 total taxa, the metric membership value will be 0.5; if there are 17 total taxa, the membership value will be 0.2; if there are 23 total taxa, the membership value will be 0.8).

BCG rules for a given level are typically comprised of multiple metrics (which are considered in combination). Together the rules for each BCG level work as a cascade from BCG level 1 to level 6, such that a sample is first tested against the level 1 rules; if the combined rule fails, then the level fails, and the assessment moves down to level 2, and so on (Figure 2). The BCG model evaluates metric membership values for all the metrics included in the rules for a given BCG level and considers the combination rules to derive the membership level for the sample. There are several different types of combination rules. If rules for two metrics are combined with an "AND" operator, then both metrics must meet the thresholds for a given BCG level (as a hypothetical example, let's say there are two rules for BCG level 3: total taxa richness ≥ 20 AND percent sensitive taxa ≥ 10%; both conditions must be met in order for the sample to be assigned to BCG level 2). If the two rules are combined with an "OR" operator (referred to as an 'alternate' rule), then *either* can be true for a sample to meet the requirements (both conditions are not necessary). Another option is having a 'best xx of xx' rule, where not all of the metrics in a group need to be met. For example, the PNMR BCG models have 'best two of three' rules, where rules for only two of the three metrics need to be met in order to meet the requirements for a given BCG level. Individual metrics that comprise the BCG rules are combined into an output that shows probability of membership in a BCG level. A sample can have full membership in a single BCG level, a tie between two levels or varying memberships among two or more levels (in which case, the level with the highest membership value is taken as the nominal level).

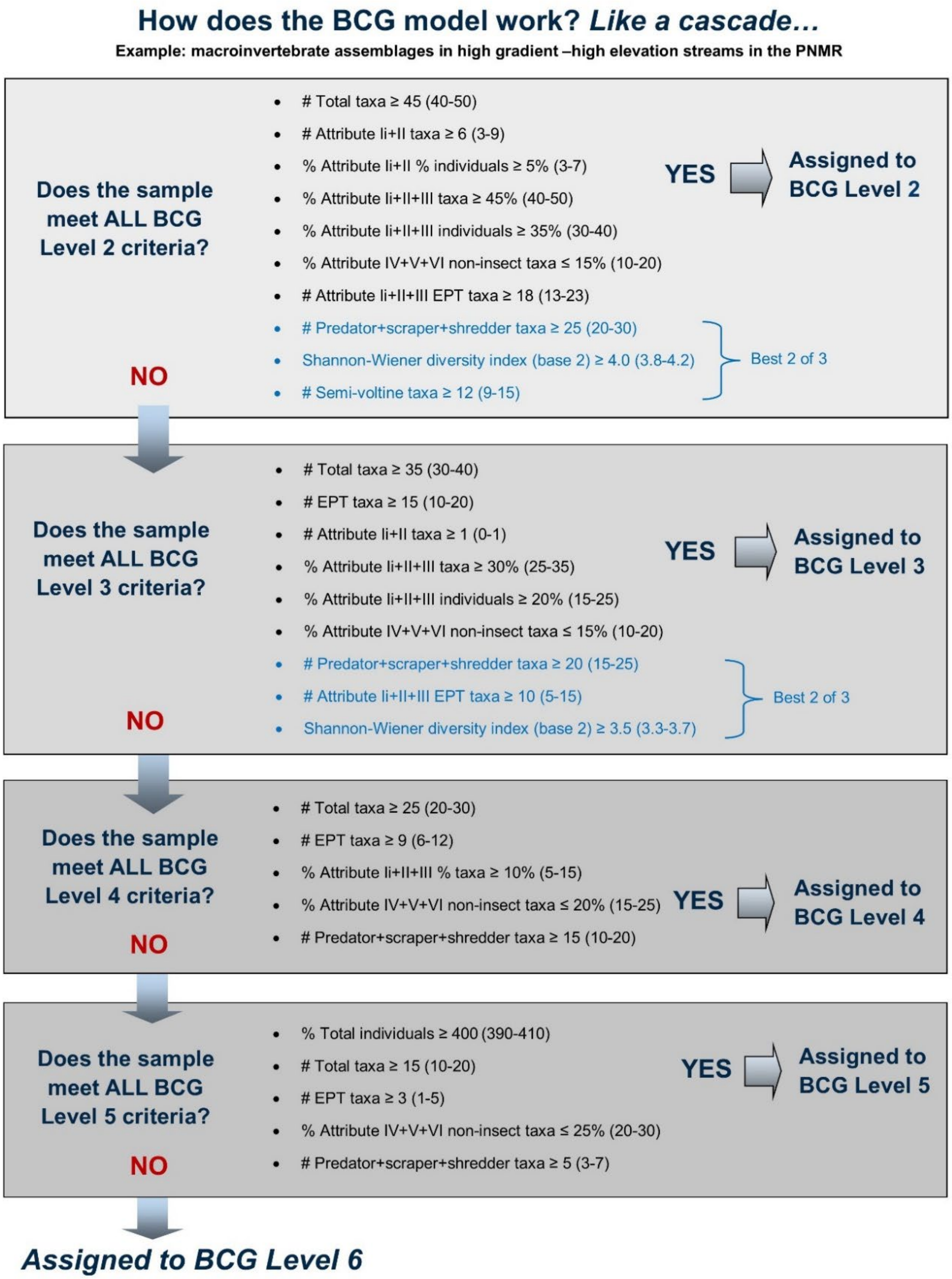

Figure 2. Example flow chart depicting how rules work as a logical cascade in the BCG model. This example is for macroinvertebrate assemblages in high gradient/high elevation freshwater wadeable PNMR streams. The flow chart starts with BCG level 2 because panelists did not assign any samples in this region to BCG level 1. Attribute Ii = sensitive, historically documented, regionally endemic; II = highly sensitive; III = intermediate sensitive; IV = intermediate tolerance; V = tolerant; VI = nonnative or intentionally introduced taxa.