Calculations for Extent, Relative Risk, and Attributable Risk: As Applied in the National Aquatic Resource Surveys

Extent Calculation

Extent calculations refer to the size, the proportion, or the percent of the resource in a condition category. Depending on the aquatic resource, the size reported may be in miles, area, or number in *Good*, *Fair*, *Poor*, or *Not Assessed* condition. While the description given focuses on stressor extent, it also applies to the biological extent.

Stressor extent (SE) in *Poor* condition is estimated as (1) the sum of the sampling weights for sites that are assessed in *Poor* condition, SE_p, (2) as the ratio of the sums of the sampling weights for the probability selected sites that are assessed in *Poor* condition divided by all sums of the sampling weights of all the selected sites regardless of condition, SEP_p or (3) the percent of stressor extent in *Poor* condition, SER_p.

SE_p is calculated as

$$SE_p = \sum_{i=1}^{n_p} w_{pi}$$

where w_{pi} is the weight for the *i*th selected site in *Poor* condition category and n_p is the number of selected sites that are in *Poor* condition.

SEP_p is calculated as

$$SEP_p = \frac{\sum_{i=1}^{n_p} w_{pi}}{\sum_{i=1}^{n} w_i}$$

where w_i is the weight for the *i*th selected site regardless of condition category and n is the total number of sites regardless of their condition category.

The stressor relative extent, SER_p, is calculated as

$$SER_p = 100 * SEP_p = 100 * \frac{\sum_{i=1}^{n_p} w_{pi}}{\sum_{i=1}^{n} w_i}$$

If the stressor extent (SE) is reported as a proportion, i.e., SEP_p, then it can be interpreted as the probability that a site chosen at random from the population will be in *Poor* stressor condition.

The total number of sites refers to all sites where the stressor is measured and a condition category is assigned, including sites that are assigned the *Not Assessed* category. Sites assigned the category *Not Assessed* may be the result of a field measurement for that indicator not being able to be measured or a field measurement was made but it was not possible to assign a condition category.

A stressor condition category may use other terminology to identify if a site is in *Poor* condition but generically we use the term *Poor*. While the focus above is on stressors, the same definitions apply to response condition variables.

Relative Risk and Attributable Risk

To estimate relative risk and attributable risk, we restrict the sites to those where both the stressor and response variable are assessed as *Good*, *Fair*, or *Poor* (or their equivalents). That is, if a site is *Not Assessed* for either the stressor or response variable, it is dropped. Next, for these sites the condition classes are combined to be either *Poor* or *Not Poor* for the stressor and response variables. For example, the *Not Poor* combines the *Good* and *Fair* condition classes. Thus, each site was designated as being in either *Poor* (P) or *Not Poor* (NP) condition, separately for each stressor and for each response variable.

The calculations that follow are based on Table 1.

Table 1: Extent estimates for response and stressor categories

Response (B)	**Stressor (S)**	
	Not Poor (NP)	Poor (P)
Not Poor (NP)	$a = \sum_{i=1}^{n_{nn}} w_{nni}$	$b = \sum_{i=1}^{n_{np}} w_{npi}$
Poor (P)	$c = \sum_{i=1}^{n_{pn}} w_{pni}$	$d = \sum_{i=1}^{n_{pp}} w_{ppi}$

In Table 1, the values in a cell are expressed in statistical terms where w_{npi} is the weight for the ith site in *Not Poor* condition for the biological category and in *Poor* condition for the stressor category, and n_{np} is the number of sites that are in *Not Poor* condition for the biological category and in *Poor* condition for the stressor category. Other cells use similar notation where, for example, *pp* stands for *Poor* response and *p* for *Poor* stressor. Note that the biological and stressor condition categories may use other terminology to identify if a site is in poor condition, but generically we use the term *Poor*. *Not Poor* is a combination of the *Good* and *Fair* categories. That is, it is all categories except *Poor* and *Not Assessed*. A separate table must be compiled for each pair of stressor and response variables.

Relative Risk Calculation

Relative risk (RR) is the ratio of the probability of a *Poor* biological condition when the stressor is *Poor* to the probability of a *Poor* biological condition when the stressor is *Not Poor*. That is,

$$RR = \frac{Pr(B = P|S = P)}{Pr(B = P|S = NP)}$$

Using the simplified notation in Table 1, relative risk (RR) is estimated as:

$$RR_{est} = \frac{d/(b + d)}{c/(a + c)}$$

A RR = 1.0 indicates there is no association between the stressor and response. That is, a *Poor* response condition in a river or stream is equally likely to occur whether or not the stressor condition is *Poor*. A RR > 1.0 indicates that a *Poor* response condition is more likely to occur when the stressor is *Poor*. For example, when the RR is 2.0, the chance that a stream is in *Poor* biological (response) condition is twice as likely when the stressor is *Poor* than when the stressor is *Not Poor*.

Further details of relative risk and its interpretation, including estimation of a confidence interval for RR_{est}, can be found in Van Sickle et al. (2006).

Attributable Risk Calculation

Population attributable risk (AR) measures what percent of the extent in *Poor* condition for a biological response variable can be attributed causally to the *Poor* condition of a specific stressor. AR is based on a scenario in which the stressor would be entirely eliminated from the aquatic resource, e.g., by means of restoration activities. That is, all the aquatic resource in *Poor* condition for the stressor are restored to the *Not Poor* condition. AR is defined as the proportional decrease in the extent of *Poor* biological response condition that would occur if the stressor were eliminated from the aquatic resource population. Mathematically, AR is defined as (Van Sickle and Paulsen 2008)

$$AR = \frac{Pr(B = P) - Pr(B = P|S = NP)}{Pr(Y = P)}$$

We estimate AR as

$$AR_{est} = \frac{BER_{est} - c/(a + c)}{BER_{est}}$$

where

$$BER_{est} = \frac{(c + d)}{(a + b + c + d)}$$

Calculation of the confidence interval for AR_{est} follows the procedure described by Van Sickle and Paulsen (2008).

A population attributable risk (AR) can take a value between 0 and 1. A value of 0 indicates either "No association" between stressor and response, or else a stressor has a zero extent, i.e., is not present in the aquatic resource population. A strict interpretation of AR in terms of stressor elimination, as described above, requires one to assume that the stressor-response relation is strongly causal and that stressor effects are reversible. Van Sickle and Paulsen (2008) discuss the reality of these assumptions, along with other issues such as interpreting them when multiple, correlated stressors are present, and using them to express the joint effects of multiple stressors.

However, AR can also be interpreted more informally, as a measure that combines RR and SE into a single index of the overall, population-level impact of a stressor on a response. Van Sickle and Paulsen (2008) show that the AR can be written as

$$AR = \frac{SER_p(RR-1)}{1+SER_p(RR-1)}$$

This shows that the numerator of AR is the product of the relative extent of a *Poor* stressor condition and the "excess" RR (i.e., RR-1) of that stressor. The denominator standardizes this product to yield AR values between 0 and 1. Thus, a high AR for a stressor indicates that the stressor is widely prevalent (has a high relative extent of *Poor* condition), and the stressor also has a large effect (high RR) in the portion of the aquatic resource where it does have *Poor* condition.

References

Van Sickle, John, and Steven P. Paulsen. 2008. "Assessing the Attributable Risks, Relative Risks, and Regional Extents of Aquatic Stressors." Journal Article. *Journal of the North American Benthological Society* 27 (4): 920–31. https://doi.org/10.1899/07-152.1.

Van Sickle, John, John L. Stoddard, Steven P. Paulsen, and Anthony R. Olsen. 2006. "Using Relative Risk to Compare the Effects of Aquatic Stressors at a Regional Scale." Journal Article. *Environmental Management* 38: 1020–30. https://doi.org/10.1007/s00267-005-0240-0.