using the index. In our data set, using the revised traits and scoring formulae in Table 14, these percentiles correspond to index values of 63 – 70 index points, respectively (Table 16). Because of the uneven distribution of reference and highly stressed index values, the percentage of highly stressed sites that are below 63 – 70 index points ranges from 91 – 96%. Using the index derived from revised traits and applying these thresholds, 91 – 96% of highly stressed sites in the current data set would be identified as biologically impacted .

One strategy for selecting a threshold is to balance errors in assessing reference and highly stressed sites: there should be as many reference sites identified as impacted as there are highly stressed sites identified as unimpacted. This is based on the premise that each data set and condition was identified with equal degrees of certainty and therefore error should be the same. Type I and Type II errors are associated with reference sites erroneously identified as impacted and highly stressed sites identified as unimpacted, respectively. In our data set, Type I and Type II errors are equal at index values at the 10th percentile, at approximately 63 index points (Table 16).

The standard deviation of the reference index distribution was 12.8 index points. A threshold of 63 index points is a little more than 1 standard deviation from the reference mean. The mean reference index score (76.4) minus 1 standard deviation is 63.7 index points.

Table 16. Low gradient IBI distribution statistics for the index calculated after trait revisions.

	All sites distribution statistics	Reference distribution statistics	Type I error	DE	Type II error
Valid N	114	27			
Minimum	7.9	34.1	0%	26.1	73.9
5th Percentile	26.2	59.5	5%	82.6	17.4
10th Percentile	33.9	63.1	10%	91.3	8.7
15th Percentile	40.5	67.1	15%	91.3	8.7
20th Percentile	43.4	69.2	20%	95.7	4.3
Lower Quartile	47.5	70.1	25%	95.7	4.3
Mean	59.9	76.4			
Median	62.4	79.1			
Upper Quartile	73.5	86.6			
Maximum	94.0	94.0			

Regression on the Calibrated Index

Similar analyses of potential thresholds were conducted using the index values derived from the calibration data; unadjusted for trait revisions. In those analyses, an index value of 60 points was the 10th percentile and balanced the Type I and Type II errors. A regression of the calibration index and the revised index showed that revised index values were generally 5 index points greater than calibration index values (Figure 27). The regression equation was $y = 0.99\ x + 4.89$ ($r^2 = 0.95$). If the regression equation is applied to the suggested calibration index threshold, the interpolated revised index threshold would be 64.3 index points.

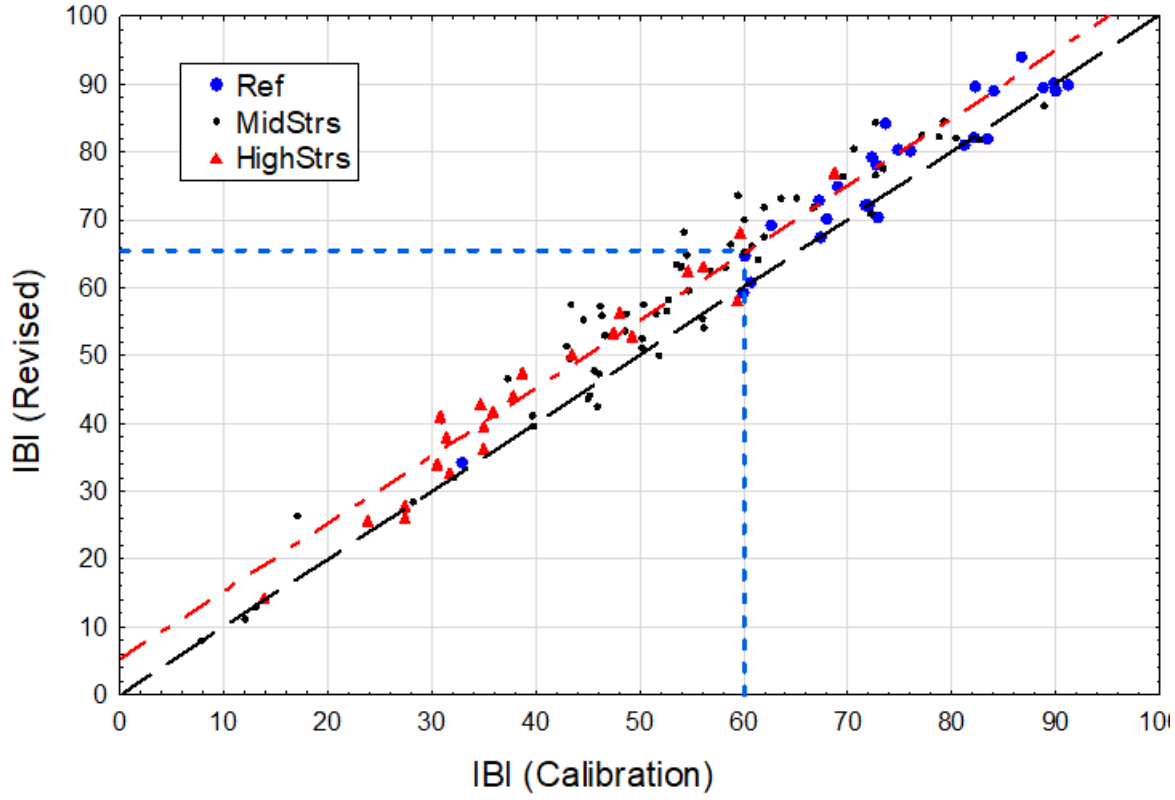

Figure 27. IBI values comparing calibration data and revised data, showing the unity line (black dashed), regression line (red dashed), and the central threshold at 60 in calibration data and 64 in revised data. The regression equation is y = 0.99 x + 4.89 (r^2 = 0.95).

These indications from reference distributions, balanced errors, standard deviations, and comparison to the preliminary calibration threshold suggest that a general condition threshold dividing satisfactory conditions from moderately degraded conditions should be in the range of 63 – 70 index points. If the balance of errors and the 10th percentile are given greater weight because they recognize potential error in both reference and highly stressed data sets and they are based on common precedent, then the threshold value would be closer to 63 index points. A general threshold of 63 index points is recommended.

Secondary Thresholds

As demonstrated in the MassDEP 100-count riffle habitat IBI threshold analyses (Stamp and Jessup 2020), secondary thresholds could be identified within the generally unimpacted and generally impacted index ranges. This would allow for refined emphasis in biological condition when prioritizing or justifying management decisions. Within the generally unimpacted index range, refined conditions could be described as Exceptional or Satisfactory based on a secondary threshold somewhat above 63 index points. A simple bisection of the unimpacted index range would suggest a threshold of 81.5 index points, half-way between the general threshold and the maximum of the index scale. In a similar fashion, the impacted range of the index scale could be bisected to describe a threshold between Moderately Degraded and Severely Degraded conditions at an index value of 31.5.

A more complex determination of secondary thresholds can be explored using proportional odds logistic regression. This technique estimates the probabilities of membership in the reference, moderately stressed, and highly stressed groups based on index values within those categories. The