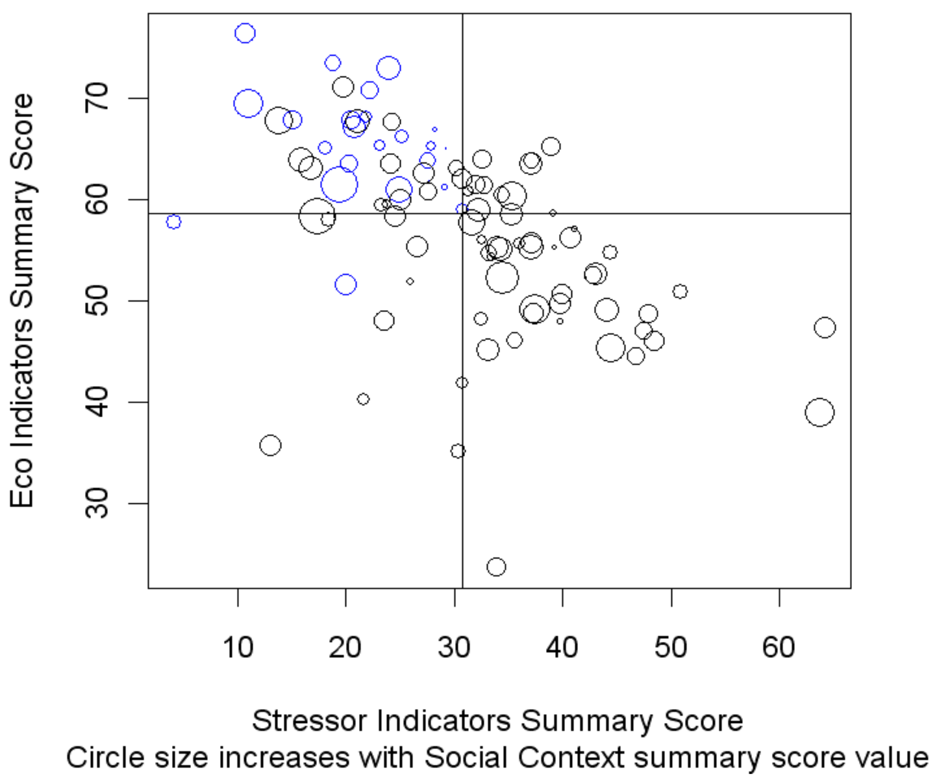
# Using Bubble Plotting as a Recovery Potential Screening Tool

Bubble plotting is a common technique for showing three or more dimensions of data in a flat, two-dimensional graph. Objects are plotted relative to conventional X and Y axes representing the first two variables, using circles (bubbles) as plotting symbols that vary in size with the value of the third variable. Assigning colors to the circles can add limited information on a fourth variable. As used in Recovery Potential Screening, bubble plotting is a way to visualize and compare multiple recovery potential scores of an entire set of watersheds simultaneously. Along with rank-ordering and mapping, bubble plotting offers a way to organize complex information about restorability, stimulate discussion and insights about differences, communicate about results and alternatives, or if desired, prescribe a clear and systematic basis for assigning priorities or decisions.

Below are techniques and a few brief examples of how bubble plotting can be used in Recovery Potential Screening. These are hypothetical examples that may use real data for demonstration purposes, but they do not constitute final analyses, policies or decisions by the US EPA or its collaborators.

Figure 1: The basic, standard recovery potential bubble plot.




**Basics of bubble plot creation**. Several commercial software products support the creation of bubble plots, but a public domain bubble plotting tool was developed specifically for Recovery Potential Screening (see [Displaying Recovery Potential Results](#) for the plotting program download and detailed directions). This tool extracts screening data from the [scoring spreadsheet](#) or similar data table format and creates a graph according to user specifications.

The standard recovery potential bubble plot (Figure 1) displays a set of waters or watersheds as individual bubbles that vary in area according to their Social Index scores, and are positioned on the graph relative to their Ecological Index score (Y axis) and Stressor Index score (X axis). Color assignment is limited but extremely useful in adding 'at-a-glance' additional information, such as identifying healthy versus impaired watersheds. Default maximum for the X and Y axes is 100, which index scores cannot exceed. Each plot also shows a horizontal and a vertical line that pass through the median Ecological and Stressor Index scores respectively, thereby separating the plot into quadrants. Users may also replace these median lines with user-selected values (e.g., statewide median values shown for reference on a bubble plot of one ecoregion's watersheds). While still in the plotting program, clicking on individual bubbles labels them with common names or IDs. Bubbles are intentionally not filled in so as not to hide smaller bubbles behind larger ones.

Although a bubble plot has a stressor axis and an ecological axis, and thus may look like a conventional 'stressor-response' relationship, this is coincidental and not what a bubble plot is meant to show. Bubble plots simply display three value gradients for the same set of objects. The basic purpose of the standard bubble plot is simply to provide an easy way to visualize an entire set of watersheds and the basic differences among them with regard to the three fundamental driving factors affecting restorability – ecological condition, stressor exposure and social context. This capability was developed because of the clear need not just for ways to estimate restorability but also for distinguishing between ecological and social factors that influence restoration success. As the examples below demonstrate, variations on bubble plotting can play numerous useful roles in Recovery Potential Screening.

## Potential Applications of Bubble Plotting

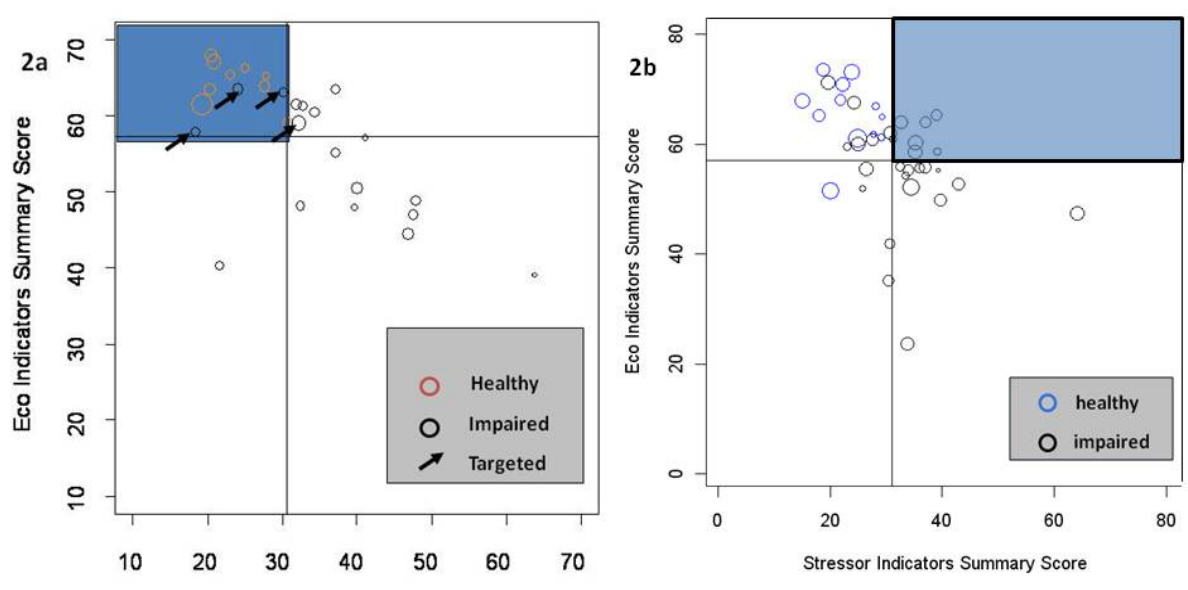
**Example 1: Compare 'healthy' and 'impaired' watersheds in general**. Although the concept of recovery potential is oriented toward restoration of impairment, most screenings include all watersheds. This presents an opportunity to visualize how known healthy watersheds score as compared with impaired watersheds in general. Just as in the identification of impairments, onsite monitoring data and a sound assessment method should be the sources of calling a watershed 'healthy'. Color-coding in Figure 1 has separated the healthy (blue) from impaired (black) watersheds in a statewide screening run. If the indicators selected for a screening are sensitive to watershed condition and restorability as intended, one would expect to see distinct pattern differences in healthy vs. impaired watersheds. In this figure, it is clear that the healthy watersheds tend toward pairing higher Ecological Index scores with lower Stressor Index scores (upper left area of the plot). Further, the distribution in the figure reveals that several of the impaired watersheds have scores similar to the healthy watersheds, whereas many others have scored much lower. Displaying the distribution of healthy and impaired together helps communicate generally that impaired watersheds may vary substantially in restorability, and some are much closer to healthy conditions than others.

When known healthy watersheds are available as reference sites, a pattern such as Figure 1 affords an important quality assurance check that the screening is sensitive to the types of overall differences among watersheds that it should be able to detect. Color-coding 'reported as impaired' from 'not reported impaired' may also be useful, but the latter should not be misinterpreted as 'healthy' because it likely combines genuinely healthy with impaired but unmonitored watersheds. Watersheds that are 'unreported' but score poorly may be candidates for improved monitoring.

**Example 2: 'Upper left quadrant' approach for targeting priority impaired watersheds**. Bubble plotting presents a logical starting point for prioritization, based on the quadrants formed by the Ecological and Stressor Index score median lines. For targeting watersheds with better restorability, the upper left quadrant of the plot normally contains the healthy watersheds, plus those watersheds that combine higher ecological and lower stressor scores yet are still impaired (see the black bubbles in Figure 2a). This simple decision rule for selecting impaired watersheds can be refined further by targeting only those upper left watersheds with higher Social Index scores as well. Yet another refinement on this approach is to generally select priorities in the upper left quadrant while allowing additional selections by expert judgment in any quadrant (see arrows in 2a). For example, a watershed in the lower left quadrant might be added to the upper left quadrant priorities because of an exceptional restoration co-funding opportunity and a very high social score.

The upper-left quadrant method for prioritizing is similar to rank-ordering by the RPI score in that it considers the influence of ecological, stressor and social factors, but it guards against underemphasizing watershed condition (i.e., the product of ecological and stressor influences). For example, it is possible that a watershed with mediocre ecological and social scores could have a high RPI score because of an abnormally low Stressor Index, and there would be little reason to target this watershed as a mediocre prospect for restoration that is under almost no threat. Such a watershed might be a priority based on rank-ordering but would not be a priority in the upper-left quadrant approach.

Figure 2: Using quadrants from bubble plots to prioritize watershed restoration. 2a (left) – upper left quadrant approach targets the most restorable watersheds; 2b (right) – upper right quadrant targets ecologically high-scoring watersheds with more severe stressors, pollutant loads or risks.

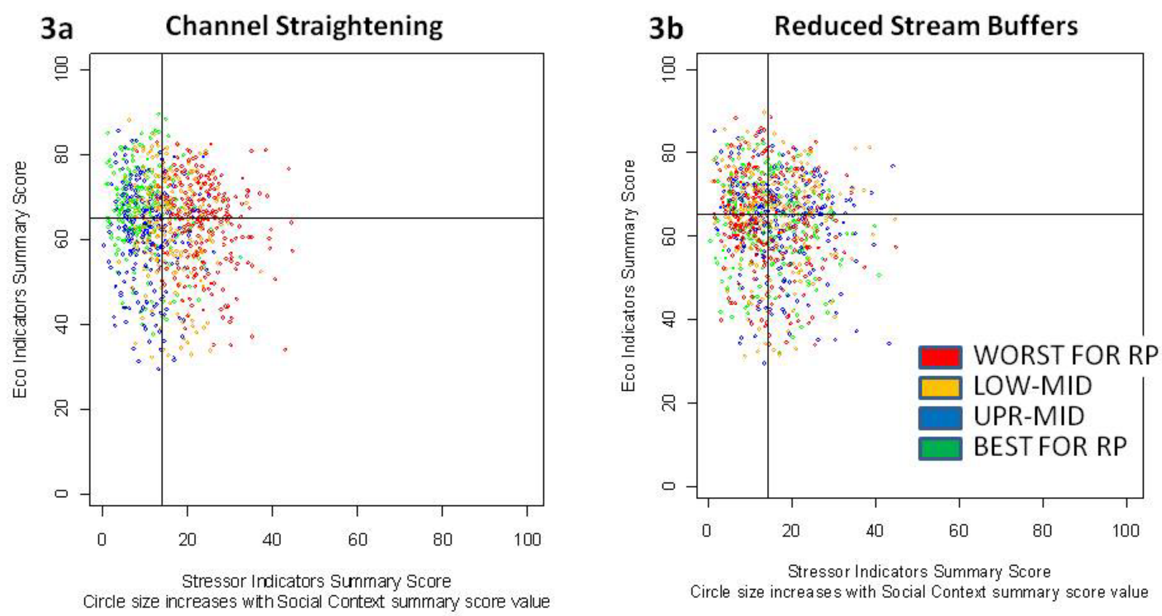



**Example 3: 'Upper right quadrant' approach for targeting priority watersheds**. A second general approach for targeting priorities is illustrated in Figure 2b. Like the previous example, this approach starts from the assumption that suitable priorities for restoring impaired watersheds are clustered, in this case in the upper right quadrant. In this example, there are no healthy watersheds in the upper right quadrant, which contains watersheds that have better than average ecological scores but also have higher stressor scores. Based on greater stressors, one might expect restoration to be more difficult in

this subset of watersheds. On the other hand, the high ecological scores suggest these are not yet severely degraded sites. Selecting priorities by using the upper right quadrangle is rooted in the idea of targeting watersheds that retain much of their functionality but are under elevated threat and/or may be providing larger pollutant loads to downstream waters. As with the upper left approach, this should be considered a starting point to which other priority considerations and refinements can be added.

**Example 4: Evaluating single indicators relative to overall patterns**. As a bubble plot is a product of many indicators together, it may raise curiosity about how closely each indicator aligns with the general screening results. Do any selected indicators run counter to the patterns observed? Which indicators appear to be the most important drivers? Again through color assignment, one can visualize how one indicator's gradient of values across all watersheds relates to the bubble plot's depiction of relative recovery potential differences overall. In Figure 3, two indicators from the same dataset are evaluated in this manner. The indicator's value range was used as the basis for assigning colors to equal-size quartiles, so that the bubble plot would show not just the overall index scores for each watershed but also its indicator-specific quartile. The channel straightening indicator's values (Figure 3a) appear to

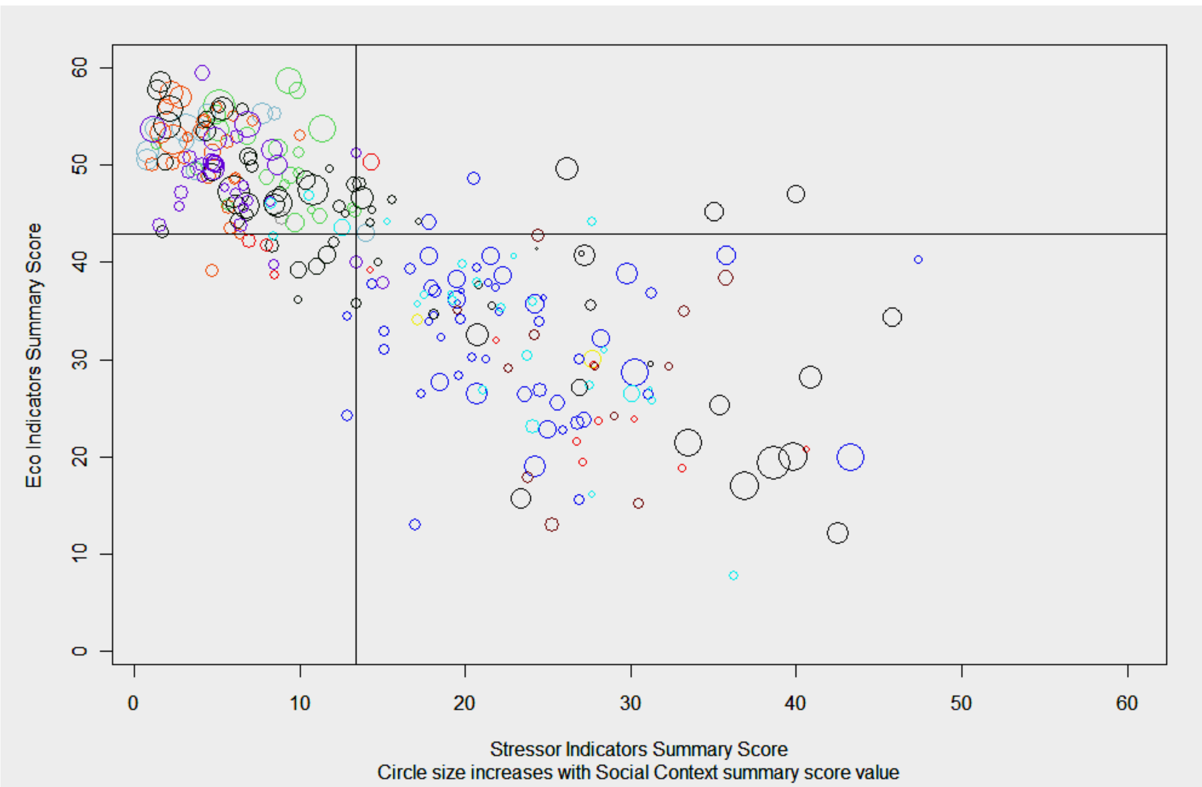
Figure 3: Screening run of over 900 small watersheds color-coded in quartiles based on two single indicator's ranges of values. The channel straightening quartiles form a pattern consistent with overall recovery potential scores, the reduced stream buffer quartiles do not.




form a distinct pattern that is consistent with the general results of the screening; in other words, watersheds that scored better for restorability on the straightening indicator (green) also tended to score well overall. In contrast, an indicator of reduced stream buffer width (Figure 3b) did not produce a discernible pattern – green is widely scattered throughout the plot, as are the other three colors. The presence or absence of pattern in this example doesn't conclusively prove that either indicator does or does not influence restorability, but it does show that one is much more consistent with the output from the full set of indicators selected for this screening run.

**Example 5: Discerning geographic zones with colors**. Another useful variation on bubble plot usage for summarizing screening information is to color-code bubbles based on location. Options include assigning specific colors to each ecoregion, major political sub-unit, or river basin. Figure 4 displays the results of a screening run color-coded to designate several major river basins; HUC10 watersheds of a given color all fall within the same basin. Several patterns are revealed in this example; for one, the river basins vary substantially in their watersheds' recovery potential. The light green, dark green and purple watersheds are consistently high-scoring, whereas the black and dark blue display elevated stressors combined with highly variable ecological condition. The red and the light blue both show wide variation in Ecological and Stressor Index scores with consistently low Social Index scores. Again in contrast, the black watersheds show some of the highest social scores despite their elevated stressors.

Figure 4: Michigan statewide watershed screening with watersheds color-coded by major river basin.




**Example 6: Compare results from independent priority ranking approaches.** Even when a bubble plot is used to display a screening-based selection of priority watersheds, it is also possible to add information for comparison of priority watersheds selected in an independent process. Figure 5 displays the results of a statewide Recovery Potential Screening of 94 watersheds. Priority watersheds identified through an independent evaluation process are color-coded and shown on the same plot, in comparison with the 'upper left quadrant' selection approach used in the screening. There was agreement on seven of ten independently selected watersheds, and two more scored just slightly outside the upper left quadrant.

Figure 5: Using a bubble plot to compare results from independent priority ranking approaches. Seven of ten priority watersheds identified through an independent evaluation process also occurred in the 'upper left quadrant' selection approach sometimes used in Recovery Potential Screening.