**Figurelegends**

Figure 1. Large feasible sets may allow better detection of deviations from the statistical baseline by generating more specific, narrowly-defined baselines. We illustrate this phenomenon using 3 hypothetical communities: a small community (S = 4, N = 34; top row), an intermediate community (S = 7, N = 71; middle row), and a large community (S= 44, N = 13360; bottom row). Larger communities have more possible SADs, and thus larger feasible sets. The large communiity has approximately 6.59e+70 possible SADs in its feasible set, while the intermediate community has 60,289 and the small community has only 297. For every SAD sampled from the feasible set (left column), we calculate the skewness (color scale) or evenness (not shown). The distributions of these values constitute the statistical baseline for comparison to observed data. We defined a “breadth index” as the ratio of the range encompassed in the one-tailed 95% density interval (space between red lines, right), compared to the full range of values for the statistic (the maximum and minimum values from the sampled feasible set). As the size of the feasible set increases and the distribution of summary statistics narrows, the breadth index decreases, indicating an increase in the specificity of the statistical baseline.

Figure 2. Many ecological communities are more skewed (top) or uneven (bottom) than their statistical baselines. Percentile ranks are calculated by comparing each community to its sampled feasible set. Very high or very low percentile ranks reflect communities that are extreme relative to their statistical baselines. The vertical red line marks the 95th percentile for skewness and the 5th percentile for evenness. At random, percentile ranks should be uniformly distributed from 0 to 100, and no more than 5% of values should be above or below the 95th and 5th percentiles, respectively. Most datasets (panels above) exhibit more communities that are highly unusual compared to their statistical baselines than would be expected by chance.

Figure 3. Feasible sets for empirical ecological communities exhibit a wide range of breadth indices (see Figure 1) for both skewness (top) and evenness (bottom). Histograms are shown for each dataset, demonstrating the range of breadth indices found for each dataset. High breadth indices indicate broad, poorly-defined statistical baselines that may impede our ability to confidently detect deviations between observations and the statistical baseline. Most datasets contain a mixture of communities ranging from broad to narrow statistical baselines, but some – particularly the skewness baseline for the Forest Inventory and Analysis – have consistently broad statistical baselines, or high values, across all of their communities. Distributions for evenness (bottom panels) appear to be more narrow than those for skewness.

Figure 4. Small communities, regardless of the dataset they come from, exhibit consistently broad statistical baselines (top), and consistently weak evidence of deviations between observed SADs and their baselines (bottom). For a subset of 371 communities from the Forest Inventory and Analysis, for which there exist communities in other datasets with matching S and N, the distributions of breadth indices (top) and percentile values (bottom) for both skewness (left) and evenness (right) do not differ between FIA (left panels) and other datasets (right panels).