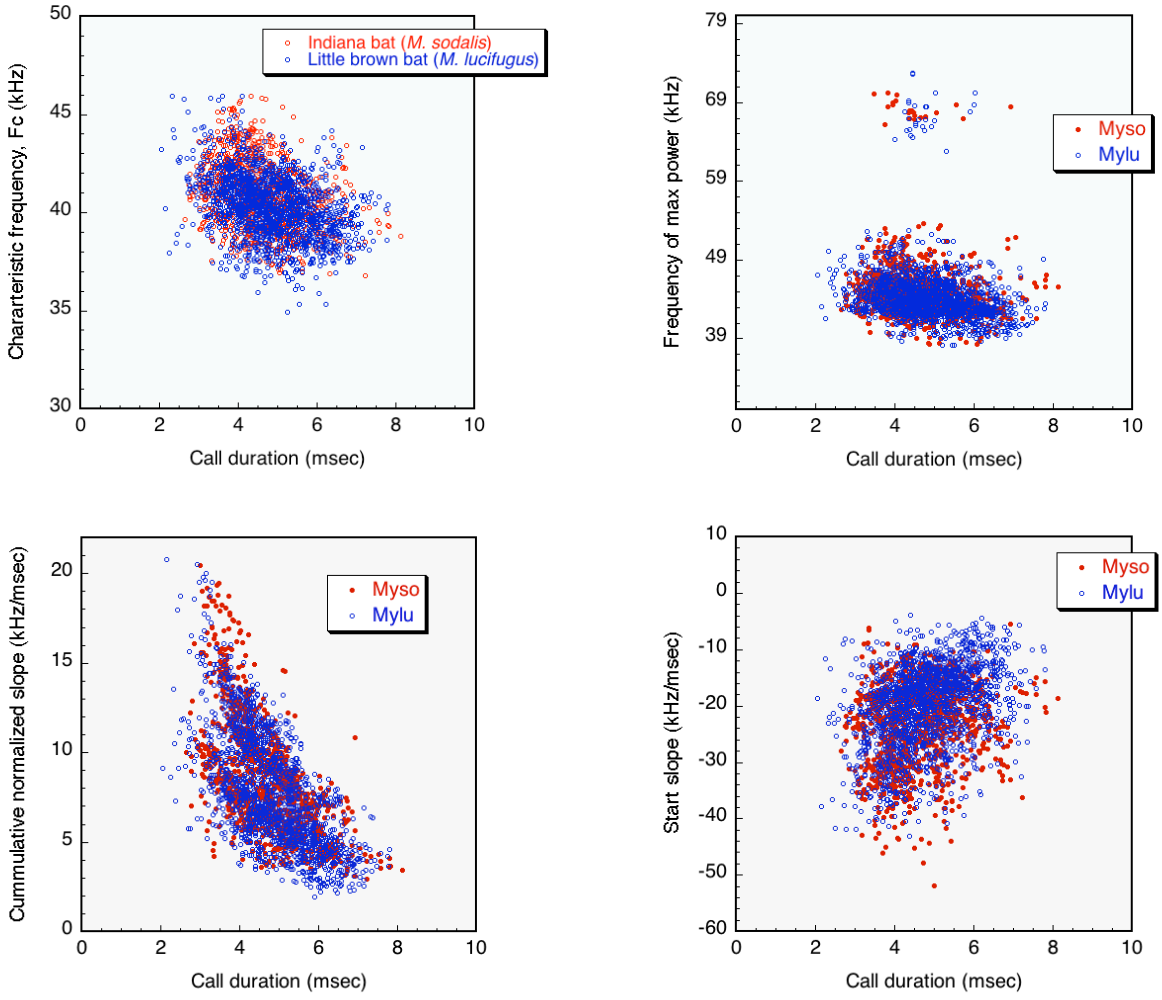
SonoBat Discrimination of Myso *vs.* Mylu

The Indiana bat, *Myotis sodalis* (Myso), and the little brown bat, *M. lucifugus* (Mylu), present substantial overlap in their echolocation call characteristics that render only a small portion of their repertoires with a tendency toward discriminating characteristics. The sample plots below display



the considerable overlap and range of call characteristics from this species pair:
Although SonoBat may report a result indicating a greater likelihood of one species over the other, e.g., 0.85 Myso versus 0.15 Mylu, such a result only indicates the relative distances from the centroid of the known multivariate data space for each species. Because these species have their centroids buried in the multivariate data clouds of the other species, they never clearly separate, and either species could just have well vocalized a call producing those results, despite lying closer to the mean values of one over the other.

To prevent outputting null species identification results, the SonoBat classifier uses this rubric: when a species decision for either of theses species does not exceed the threshold discriminant probability setting (DP, SonoBat uses 0.90 as the default setting), and if the second potential species comes out as the opposite of this pair, and their combined discriminant probability score meets or exceeds the threshold setting, then SonoBat will output this result using the ambiguous designation “MysoMylu.” This will indicate the call or sequence probably came from one of these two species, but presented call characteristics within overlapping data space that prevented disambiguation.

However, with most Myotis species, ***longer duration calls provide more robust and consistent data that enhances discrimination performance***. The following results parse classification results by call duration for 373 *M. sodalis* and *M. lucifugus* sequences recorded in IN, IL, MO, KY, TN, PA, NJ, and VT that yielded 2,735 parameterized calls using a maximum of 8 calls to consider per file, and a quality acceptance threshold of 0.80[1]:

Individual call results with discriminant probability threshold of **0.90**.

call duration (msec)		%correct	%accepted[2]
↓	Myso	33.3	*0.7*
↓	Mylu	54.5	*1.7*
4.0	Myso/Mylu[3]	100.0	*34.9*
4.0	Myso	91.7	*5.2*
↓	Mylu	85.4	*6.4*
5.0	**Myso/Mylu[3]**	**100.0**	***37.4***
5.0	Myso	92.7	*28.7*
↓	Mylu	92.1	*25.3*
6.5	**Myso/Mylu[3]**	**100.0**	***41.1***
6.5	**Myso**	**100.0**	***54.3***
↓	**Mylu**	**100.0**	***45.3***
	Myso/Mylu[3]	**100.0**	***35.4***

Generally, calls less than 5.0 msec yield unreliable classification. Note the very low %acceptance and ambiguous classification for calls less than 4.0 msec. Although calls between 4.0 and 5.0 msec achieved reasonable %correct classification, the low %accepted for these calls may reflect an artifact from the classifier being based on a finite data set rather than absolute performance on an open data set from the wild. Accepting less than ~25% of the sample indicates a weak, non-robust discrimination that will likely produce unreliable results with actual field data, i.e., the inherent nature of the call characteristics do not separate well for confident discrimination.

As a first rule of thumb, have little confidence in any results from calls less than 4.0 msec, begin to have some confidence in calls from 4.0–5.0 msec, and have increasing confidence for calls longer than that. And of course, have more confidence in sequence results (based on longer calls) than from individual calls. The sequence decisions that SonoBat generates integrate the combined information from all calls in a sequence (see below).

SonoBat classifies calls and sequences using an expert system incorporating an ensemble consensus of redundant hierarchical decision algorithms and reports a single species decision when that result exceeds the discriminant probability (DP) threshold at each decision, *and* passes post-decision checks of known call characteristics. This decision path optimization approach outperformed tests using other standard machine intelligence systems (Artificial Neural Networks,

[1] ***The results reported here represent idealized classification performance based on high quality recordings*** made with Pettersson D240X and D500X detectors, and with Binary Acoustic Technology AR125 detectors. ***Actual performance will decline commensurate with recording quality*** (see *Recommendations for quality recording* in the NE Classification notes document).

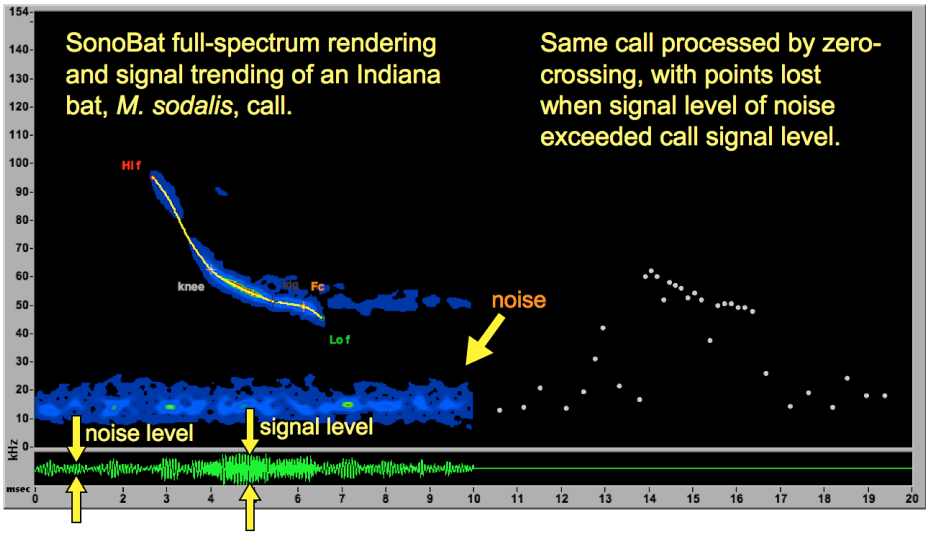
[2] Values listed as %correct considered just those results that emerged from the classifier at or above a discriminant probability threshold of 0.90. The %accepted reports the proportion of the sample that met or exceeded the discriminant probability threshold, whether correct or incorrect.

[3] Myso/Mylu indicates a result of MysoMylu, Myso, or Mylu, whether correct or incorrect for Myso (if Mylu) or Mylu (if Myso), i.e., the overall rate for correctly discriminating this species pair from other species.

Bayesian, etc.). SonoBat reports the DP of the final decision step, in the cases above, that from the decision discriminating between *M. sodalis* and *M. lucifugus*.

Mean call duration (msec)		Sequence results by mean call duration (%correct *%accepted*)								
↓	Myso	by vote:	88.9	*9.6*	mean sqnc:	100.0	*3.6*	agreement:	100.0	*2.4*
5.0	Mylu	by vote:	100.0	*11.1*	mean sqnc:	75.0	*7.1*	agreement:	100.0	*1.6*
5.0-	Myso	by vote:	93.8	*38.5*	mean sqnc:	84.6	*28.2*	agreement:	88.2	*38.5*
-6.0	Mylu	by vote:	100.0	*38.2*	mean sqnc:	94.4	*25.0*	agreement:	96.7	*42.6*
6.0	Myso	by vote:	100.0	*88.9*	mean sqnc:	100.0	*88.9*	agreement:	100.0	*88.9*
↓	Mylu	by vote:	100.0	*85.2*	mean sqnc:	100.0	*70.4*	agreement:	100.0	*85.2*

In the penultimate hierarchical decision before attempting to discriminate between *M. sodalis* and *M. lucifugus*, SonoBat first discriminates this similar species pair, i.e., MysoMylu, from other Myotis species. SonoBat discriminates the composite MysoMylu with greater performance than with the individual species. MysoMylu decisions that meet or exceed the DP threshold will then be passed on to the classifier decision steps that attempt to discriminate between *M. sodalis* and *M. lucifugus*. Calls that reach this step in the decision hierarchy that do not meet or exceed the DP threshold output as MysoMylu by the rubric described above.



To minimize misclassifications, SonoBat performs quality control by assessing a number of signal quality and reliability indicators. If a call fails accepted thresholds for any of these indicators, SonoBat rejects the result from automated classification as it can indicate a poor quality signal that can lead to misclassification. (During manual inspection, SonoBat will report the classification result of calls that fail any of the reliability tests but will gray out the display to indicate an unreliable result.)

The amplitude and multiple frequency content of full-spectrum data enable assessment of signal quality and evaluation of the acoustic environment of the recording site. For example, one such measure, the signal to noise ratio (SNR), measures the relative strength of a signal of interest (the call) to the strength of the background signal level (a measure unavailable in zero-crossing data that can only access the dominant frequency at any time). Calls with low SNR more often render poor data that can lead to misclassification. Multiple frequency content also enables more reliable tracking of call trends, and better call data extraction, through echo clutter and ambient noise (see figure). For more information, see

http://www.sonobat.com/download/FullSpect_and_Zero-Crossing.ppt

Reliable acoustic species classification depends strongly on high quality, distortion-free recordings. For recommendations on recording and further information and guidance on using and interpreting the SonoBat classifier, see

http://www.sonobat.com/download/SonoBatClassificationNote-NE-v306.pdf