United States Department of The Interior
U.S. Geological Survey
Ecosystems Division

Virginia Cooperative Fish and Wildlife Research Unit   Fax: (540) 231-7580
106 Cheatham Hall   Phone: (540) 231-5927
Virginia Tech   Email: wmford@vt.edu
Blacksburg, VA 24061

September 15, 2014

Memo to: Mike Armstrong, Andrew King and Robin Niver

From: Mark Ford

Subject: Echolocation Identification Software Results

Methods

Over the summer, the U.S. Geological Survey Virginia Cooperative Fish and Wildlife Research Unit tested Kaleidoscope 2.0.3, BCID 2.6a, and Echoclass 3 per the request of the U.S. Fish and Wildlife Service for the ongoing development of new survey protocols for the Indiana bat (*Myotis sodalis*) and other species. Using echolocation pulses from hand-release and manually identified wild pulses from a variety of sources/data sets from Dr. Lynn Robbins, Dr. Kevin Murray, Dr. Eric Britzke, Mr. Santiago Martin, Mr. Chris Dobony and myself, I created several "test" groupings of species assemblages based on bat community membership and activity proportions from northwestern New York (Ford et al. 2011: Journal of Fish and Wildlife Management 2:125-134), north-central West Virginia (Ford et al. 2005: Biological Conservation 126:528-539), northwest Georgia (Johnson et al. 2010: Folia Zoologica 59:192-202), south-central South Carolina (Ford et al. 2006: Journal of Wildlife Management 70:1174-1184) and southern Missouri (Murray et al. 1999: Acta Chiropterologica 105:105-112; Amelon 2007; PhD Dissert. Univ. Missouri). Datasets also had "double-calls" inserted and other aberrant pulses, i.e., northern flying squirrel (*Glaucomys sabrinus*). With the help of Mr. Alex Silvis, a PhD student in the Department of Fish and Wildlife Conservation at Virginia Tech, we ran each set through the various software programs. Alex then collated the output (for p-values, lower is better; for other measures higher is better) and also performed *post-hoc* assessments of overall correct classification, sensitivity (species correctly identified when present, a measure of true positive) and specificity (species not identified when absent, a measure of true negative) as well as threshold adjusted measures of species proportion, area under the curve of the receiver operating characteristic, omission and per-species correct classification.

We evaluated program predictive performance for file level assignment of species identity using sensitivity, specificity, and the area under the curve of the receiver operating characteristic. We calculated these metrics in R (v. 3.0.2) using library SDMTools (VanDerWal et al. 2012; R package version 1.1-13; http://CRAN.R-project.org/package=SDMTools) after converting multi-species output to binary classification (i.e., *Myotis sodalis* versus not *Myotis sodalis*). We set the threshold value for each species to be equal to its proportional abundance in the test call library. For the New York and West Virginia datasets, runs were performed with northern long-eared bats (*Myotis septentrionalis*) removed and Indiana bats retained as well as the converse to further assess sensitivity and specificity. For each dataset, results are reported for overall software

identification by species and program assignment of confidence (p-value), each program's identification with *a priori* call assignments and overall correct classification, the software's full classification metrics and finally individual call identification.

Results

Data are reported from simulated New York (Tables 1-7), West Virginia (Tables 8-15), Georgia (Tables 16-23) and South Carolina (Tables 24-31) software runs. The Missouri dataset contained some corrupted files and final analyses have not been completed; once the corrupted files are identified, we will attempt analyses of these data later in the fall.   Kaleidoscope, BCID and Echoclass each passed the classification and p-value standards set by the U.S. Fish and Wildlife Service for Indiana bats and gray bats (*Myotis grisescens*) when examining mixed species sets of echolocation data. Programs were not penalized for unknown or no identifications in the *post-hoc* assessment. Our initial run of Echoclass was complicated by the program's grouping of echolocation pulses by actual time stamp rather than out folder arrangement. While this is a fail-safe mechanism to prevent errant or unscrupulous grouping of passes from different sites or nights, we had to request a slightly modified version of Echoclass that would permit batch identification of our simulated data that was comprised of multiple echolocation passes from a variety of sites and across a several year period. The resulting reanalysis of Echoclass provided *post-hoc* assessment of confidence more in line with the actual program's per file identification.

Per software reporting guidelines, both BCID and Echoclass had confidence in correctly identifying northern long-eared bats. Kaleidoscope showed relatively high sensitivity in the *post-hoc* assessment and in the file by file identification, however, it showed little overall confidence in confirming northern long-eared bat presence when it fact the species was in the datasets. For small-footed bats (*Myotis leibii*), both Kaleidoscope and BCID performed well, whereas Echoclass had little confidence in their presence and did not perform well in *post-hoc* assessments. The converse was true for the Southeastern bat (*Myotis austroriparius*), as Echoclass was the only software program that identified the species with confidence. It should be noted that the Kaleidoscope version tested does not provide Southeastern bat in its species pool. For all the other remaining bat species, the software programs performed relatively well.

When datasets were run with no Indiana bats (Table 29) but were allowed to consider northern long-eared bats, all programs falsely identified presence ($P < 0.03$) of Indiana bats. Both BCID and Echoclass returned high confidence for presence of northern long-eared bats in this situation, whereas Kaleidoscope returned a false negative. When datasets were run with no northern long-eared bats but were allowed to consider Indiana bats, Kaleidoscope and Echoclass correctly determined absence for both test datasets, but BCID incorrectly documented presence in one of our test datasets. All three programs were able to correctly identify presence of Indiana bat calls in the absence of northern long-eared bat calls.

Recommendations

At the individual file level, each software program clearly made some misidentifications but within the parameters set for use, each should be acceptable for Indiana bats, gray bats, northern long-eared bats, little brown bats (*Myotis lucifugus*) and the non-myotids. It appears from liberal to conservative, the software should be ranked BCID, Kaleidoscope and Echoclass. Suggestions for improvement for BCID would be to continue to refine the program to produce higher degrees of specificity for Indiana bat and northern long-eared bat. Kaleidoscope needs to incorporate Southeastern bat into its configuration and improve sensitivity for northern long-eared bats. Similarly Echoclass needs to improve sensitivity for small-footed bats. That being said, in settings where Southeastern bats could occur, Echolass at present, would the most acceptable program.

Table 1. Post-identification per species confidence from maximum likelihood estimator or surrogate (see program specifications for details) for known echolocation pulses for Kaleidoscope, BCID and Echoclass for simulated New York dataset based on Ford et al. (2011). Note: programs may identify echolocation pulses to a species without high confidence (> 0.1-0).

_____

Kaleidoscope

ID	N	p
EPFU	17	0.003225
LABO	26	0.000126
LACI	21	0
LANO	21	0.000016
MYLE	13	0
MYLU	82	0
MYSE	12	1
MYSO	16	0.000268
PESU	10	3.6E-06
NO ID	23	
NOISE	2	

BCID

ID	N	p
EPFU	21	0.000001
LABO	25	0.000001
LACI	22	0.000001
LANO	10	0.000001
MYLE	8	0.000001
MYLU	79	0.000001
MYSE	12	0.000001
MYSO	41	0.000001
PESU	13	0.000001
UNKN	4	

Echoclass

ID	N	p
EPFU	16	0
LABO	46	0
LACI	17	0.0322
LANO	5	1
MYLE	4	1
MYLU	30	0
MYSE	5	0
MYSO	37	0
PESU	10	0
UNKN	68	

_____

Table 2. Per species echolocation pulse assignment and overall correct classification rate for Kaleidoscope for simulated New York dataset based on Ford et al. (2011). Columns in bold are known echolocation pulse assignments. Rows in italics are Kaleidoscope assignment.

	EPFU	**LABO**	**LACI**	**LANO**	**MYLE**	**MYLU**	**MYSE**	**MYSO**	**MISC/ODD**	**PESU**
EPFU	16	0	0	1	0	0	0	0	0	0
LABO	0	19	0	0	0	6	0	0	1	0
LACI	0	0	21	0	0	0	0	0	0	0
LANO	10	0	1	9	0	0	0	0	1	0
MYLE	0	0	0	0	6	0	2	5	0	0
MYLU	0	0	0	0	1	69	0	9	3	0
MYSE	0	0	0	0	0	2	10	0	0	0
MYSO	0	0	0	0	2	5	3	6	0	0
UNKN	0	3	0	0	1	9	3	7	0	0
PESU	0	0	0	0	0	0	0	0	0	10

%
correct   77.93427

Table 3. Per species echolocation pulse assignment performance for Kaleidoscope for simulated New York dataset based on Ford et al. (2011). Threshold is the proportional abundance of per species echolocation pulses. AUC is area under the curve. Omission is proportion of occurrences misidentified adjusted by the threshold value. Sensitivity is the proportion of true events assigned correctly (assessment of true positive). Specificity is the proportion of non-events assigned correctly (assessment of true negative). Correct is per species overall correct classification. Kappa measures the proportion of correct classification after accounting for the probability of chance agreement.

	Threshold	AUC	Omission	Sensitivity	Specificity	Correct	Kappa
EPFU	0.119266	0.805088	0.384615	0.615385	0.994792	0.949541	0.71755
LABO	0.087156	0.982412	0	1	0.964824	0.96789	0.827023
LACI	0.100917	0.977273	0.045455	0.954545	1	0.995413	0.974201
LANO	0.045872	0.921154	0.1	0.9	0.942308	0.940367	0.552856
MYLE	0.041284	0.816587	0.333333	0.666667	0.966507	0.954128	0.522139
MYLU	0.376147	0.872938	0.158537	0.841463	0.904412	0.880734	0.745875
MYSE	0.068807	0.828407	0.333333	0.666667	0.990148	0.96789	0.723851
MYSO	0.091743	0.624747	0.7	0.3	0.949495	0.889908	0.27414
PESU	0.045872	1	0	1	1	1	1

Table 4. Per species echolocation pulse assignment and overall correct classification rate for BCID for simulated New York dataset based on Ford et al. (2011). Columns in bold are known echolocation pulse assignments. Rows in italics are BCID assignment.

	EPFU	**LABO**	**LACI**	**LANO**	**MYLE**	**MYLU**	**MYSE**	**MYSO**	**PESU**	**MISC/ODD**
EPFU	19	0	1	2	0	0	0	0	0	0
LABO	0	19	0	0	0	5	0	0	1	1
LACI	5	0	16	1	0	0	0	0	0	0
LANO	2	0	1	7	0	0	0	0	0	0
MYLE	0	0	0	0	9	0	0	0	0	0
MYLU	0	0	0	0	0	73	0	4	0	2
MYSE	0	0	0	0	1	0	12	0	0	0
MYSO	0	0	0	0	1	12	4	24	0	1
PESU	0	4	0	0	0	0	0	0	10	0
UNKN	0	0	1	0	0	2	1	0	0	0

%
correct   81.11588

Table 5. Per species echolocation pulse assignment performance for BCID for simulated New York dataset based on Ford et al. (2011). Threshold is the proportional abundance of per species echolocation pulses. AUC is area under the curve. Omission is proportion of occurrences misidentified adjusted by the threshold value. Sensitivity is the proportion of true events assigned correctly (assessment of true positive). Specificity is the proportion of non-events assigned correctly (assessment of true negative). Correct is per species overall correct classification. Kappa measures the proportion of correct classification after accounting for the probability of chance agreement.

_____

		Threshold	AUC	Omission	Sensitivity	Specificity	Correct	Kappa
1	EPFU	0.109705	0.858276	0.269231	0.730769	0.985782	0.957806	0.768374
2	LABO	0.097046	0.896688	0.173913	0.826087	0.96729	0.953586	0.749736
3	LACI	0.075949	0.930746	0.111111	0.888889	0.972603	0.966245	0.781768
4	LANO	0.042194	0.843392	0.3	0.7	0.986784	0.974684	0.686784
5	MYLE	0.046414	0.909091	0.181818	0.818182	1	0.991561	0.895641
6	MYLU	0.379747	0.885147	0.188889	0.811111	0.959184	0.902954	0.788991
7	MYSE	0.067511	0.872738	0.25	0.75	0.995475	0.978903	0.816478
8	MYSO	0.118143	0.885509	0.142857	0.857143	0.913876	0.907173	0.633797
9	PESU	0.046414	0.945696	0.090909	0.909091	0.982301	0.978903	0.789033

_____

Table 6. Per species echolocation pulse assignment and overall correct classification rate for Echoclass for simulated New York dataset based on Ford et al. (2011). Columns in bold are known echolocation pulse assignments. Rows in italics are Echoclass assignment.

	EPFU	**LABO**	**LACI**	**LANO**	**MYLE**	**MYLU**	**MYSE**	**MYSO**	**PESU**	**MISC/ODD**
EPFU	15	0	0	1	0	0	0	0	0	0
LABO	0	20	0	0	0	18	2	4	1	1
LACI	0	0	17	0	0	0	0	0	0	0
LANO	2	0	0	3	0	0	0	0	0	0
MYLE	0	0	0	0	3	0	0	1	0	0
MYLU	0	0	0	0	0	28	0	1	0	1
MYSE	0	0	0	0	1	0	4	0	0	0
MYSO	0	0	0	0	0	16	4	17	0	0
NOISE	0	0	0	0	1	0	4	0	0	0
PESU	0	0	0	1	0	1	0	0	8	0
UNKN	9	2	5	5	5	28	6	4	1	3

%
correct   68.45238

Table 7. Per species echolocation pulse assignment performance for Echoclass for simulated New York dataset based on Ford et al. (2011). Threshold is the proportional abundance of per species echolocation pulses. AUC is area under the curve. Omission is proportion of occurrences misidentified adjusted by the threshold value. Sensitivity is the proportion of true events assigned correctly (assessment of true positive). Specificity is the proportion of non-events assigned correctly (assessment of true negative). Correct is per species overall correct classification. Kappa measures the proportion of correct classification after accounting for the probability of chance agreement.

_____

		Threshold	AUC	Omission	Sensitivity	Specificity	Correct	Kappa
1	EPFU	0.1	0.937908	0.117647	0.882353	0.993464	0.982353	0.899329
2	LABO	0.117647	0.913333	0	1	0.826667	0.847059	0.528785
3	LACI	0.1	1	0	1	1	1	1
4	LANO	0.029412	0.793939	0.4	0.6	0.987879	0.976471	0.587879
5	MYLE	0.023529	0.871988	0.25	0.75	0.993976	0.988235	0.743976
6	MYLU	0.370588	0.712876	0.555556	0.444444	0.981308	0.782353	0.47714
7	MYSE	0.058824	0.696875	0.6	0.4	0.99375	0.958824	0.514286
8	MYSO	0.135294	0.801538	0.26087	0.73913	0.863946	0.847059	0.479878
9	PESU	0.052941	0.938233	0.111111	0.888889	0.987578	0.982353	0.832787

_____

Table 8. Post-identification per species confidence from maximum likelihood estimator or surrogate (see program specifications for details) for known echolocation pulses for Kaleidoscope, BCID and Echoclass for simulated West Virginia dataset based on Ford et al. (2005). Note: programs may identify echolocation pulses to a species without high confidence (> 0.1-0).

Kaleidoscope			BCID			Echoclass		
ID	N	p	ID	N	p	ID	N	p
EPFU	17	1	EPFU	24	0.000001	EPFU	16	0
LABO	41	0	LABO	42	0.000001	LABO	53	0
LACI	21	0.000043	LACI	21	0.000001	LACI	18	0.0192
LANO	15	0.003939	LANO	6	0.000099	LANO	5	1
MYLE	10	0.030885	MYLE	3	0.039812	MYLE	3	1
MYLU	51	0	MYLU	53	0.000001	MYLU	25	0
MYSE	19	1	MYSE	18	0.000001	MYSE	3	0
MYSO	21	0.000972	MYSO	39	0.000001	MYSO	35	0
PESU	19	0	PESU	17	0.000001	PESU	8	0
NO ID	19		UNKN	2		UNKN	67	

Table 9. Per species echolocation pulse assignment and overall correct classification rate for Kaleidoscope for simulated West Virginia dataset based on Ford et al. (2005). Columns in bold are known echolocation pulse assignments. Rows in italics are Kaleidoscope assignment.

	EPFU	**LABO**	**LACI**	**LANO**	**MYLE**	**MYLU**	**MYSE**	**MYSO**	**MISC/ODD**	**PESU**
EPFU	16	0	0	1	0	0	0	0	0	0
LABO	0	36	0	0	0	4	0	0	0	1
LACI	0	0	21	0	0	0	0	0	0	0
LANO	10	0	1	4	0	0	0	0	0	0
MYLE	0	0	0	0	4	0	1	5	0	0
MYLU	0	0	0	0	0	43	0	5	3	0
MYSE	0	0	0	0	1	2	16	0	0	0
MYSO	0	0	0	0	2	5	4	10	0	0
NO ID	0	3	0	0	0	6	3	7	0	0
PESU	0	1	0	0	2	0	0	0	2	14

%
correct   78.4689

Table 10. Per species echolocation pulse assignment performance for Kaleidoscope for simulated West Virginia dataset based on Ford et al. (2005). Threshold is the proportional abundance of per species echolocation pulses. AUC is area under the curve. Omission is proportion of occurrences misidentified adjusted by the threshold value. Sensitivity is the proportion of true events assigned correctly (assessment of true positive). Specificity is the proportion of non-events assigned correctly (assessment of true negative). Correct is per species overall correct classification. Kappa measures the proportion of correct classification after accounting for the probability of chance agreement.

	Threshold	AUC	Omission	Sensitivity	Specificity	Correct	Kappa
EPFU	0.121495	0.805033	0.384615	0.615385	0.994681	0.948598	0.716999
LABO	0.172897	0.972362	0.027027	0.972973	0.971751	0.971963	0.905989
LACI	0.102804	0.977273	0.045455	0.954545	1	0.995327	0.974148
LANO	0.023364	0.873684	0.2	0.8	0.947368	0.943925	0.378208
MYLE	0.042056	0.707588	0.555556	0.444444	0.970732	0.948598	0.394236
MYLU	0.252336	0.873148	0.203704	0.796296	0.95	0.911215	0.760288
MYSE	0.098131	0.87318	0.238095	0.761905	0.984456	0.962617	0.779438
MYSO	0.093458	0.721649	0.5	0.5	0.943299	0.901869	0.433577
PESU	0.070093	0.954104	0.066667	0.933333	0.974874	0.971963	0.80853

Table 11. Per species echolocation pulse assignment and overall correct classification rate for BCID for simulated West Virginia dataset based on Ford et al. (2006). Columns in bold are known echolocation pulse assignments. Rows in italics are BCID assignment.

	EPFU	**LABO**	**LACI**	**LANO**	**MYLE**	**MYLU**	**MYSE**	**MYSO**	**PESU**	**MISC/ODD**
EPFU	19	0	1	3	0	0	0	0	0	1
LABO	0	34	0	0	1	3	0	0	3	1
LACI	5	0	16	0	0	0	0	0	0	0
LANO	2	1	1	2	0	0	0	0	0	0
MYLE	0	0	0	0	3	0	0	0	0	0
MYLU	0	0	0	0	1	51	0	0	0	2
MYSE	0	0	0	0	1	0	17	1	0	0
MYSO	0	0	0	0	4	7	1	26	0	1
PESU	0	5	0	0	0	0	0	0	13	0
UNKN	0	1	1	0	0	0	1	0	0	0

%
correct   81.90045

Table 12. Per species echolocation pulse assignment performance for BCID for simulated West Virginia dataset based on Ford et al. (2005). Threshold is the proportional abundance of per species echolocation pulses. AUC is area under the curve. Omission is proportion of occurrences misidentified adjusted by the threshold value. Sensitivity is the proportion of true events assigned correctly (assessment of true positive). Specificity is the proportion of non-events assigned correctly (assessment of true negative). Correct is per species overall correct classification. Kappa measures the proportion of correct classification after accounting for the probability of chance agreement.

_____

Species	threshold	AUC	Omission	Sensitivity	Specificity	Correct	Kappa
EPFU	0.115044	0.852885	0.269231	0.730769	0.975	0.946903	0.730203
LABO	0.176991	0.903495	0.15	0.85	0.956989	0.938053	0.791458
LACI	0.079646	0.932425	0.111111	0.888889	0.975962	0.969027	0.803673
LANO	0.022124	0.69095	0.6	0.4	0.9819	0.969027	0.347898
MYLE	0.044248	0.65	0.7	0.3	1	0.969027	0.450313
MYLU	0.269912	0.908942	0.163934	0.836066	0.981818	0.942478	0.848572
MYSE	0.079646	0.967415	0.055556	0.944444	0.990385	0.986726	0.911696
MYSO	0.119469	0.948818	0.037037	0.962963	0.934673	0.938053	0.753005
PESU	0.070796	0.894345	0.1875	0.8125	0.97619	0.964602	0.745639

_____

Table 13. Per species echolocation pulse assignment and overall correct classification rate for Echoclass for simulated West Virginia dataset based on Ford et al. (2005). Columns in bold are known echolocation pulse assignments. Rows in italics are Echoclass assignment.

	EPFU	LABO	LACI	LANO	MYLE	MYLU	MYSE	MYSO	PESU	MISC/ODD
EPFU	15	0	0	1	0	0	0	0	0	0
LABO	0	36	0	0	0	9	4	1	2	1
LACI	0	0	17	1	0	0	0	0	0	0
LANO	2	0	0	2	0	0	0	0	0	1
MYLE	0	0	0	0	0	0	1	2	0	0
MYLU	0	0	0	0	0	24	0	0	0	1
MYSE	0	0	0	0	0	0	3	0	0	0
MYSO	0	0	0	0	1	12	2	20	0	0
NOISE	0	0	0	0	1	0	5	0	0	0
PESU	0	0	0	0	0	0	0	0	7	1
UNKN	9	4	5	1	8	15	13	4	6	2

%
correct   76.54321

Table 14. Per species echolocation pulse assignment performance for Echoclass for simulated West Virginia dataset based on Ford et al. (2005). Threshold is the proportional abundance of per species echolocation pulses. AUC is area under the curve. Omission is proportion of occurrences misidentified adjusted by the threshold value. Sensitivity is the proportion of true events assigned correctly (assessment of true positive). Specificity is the proportion of non-events assigned correctly (assessment of true negative). Correct is per species overall correct classification. Kappa measures the proportion of correct classification after accounting for the probability of chance agreement

_____

	Threshold	AUC	Omission	Sensitivity	Specificity	Correct	Kappa
EPFU	0.10241	0.937821	0.117647	0.882353	0.993289	0.981928	0.899068
LABO	0.216867	0.934615	0	1	0.869231	0.89759	0.742471
LACI	0.10241	0.996644	0	1	0.993289	0.993976	0.968065
LANO	0.024096	0.740741	0.5	0.5	0.981481	0.96988	0.429161
MYLE	0.006024	0.509091	1	0	0.981818	0.975904	-0.00912
MYLU	0.271084	0.762534	0.466667	0.533333	0.991736	0.86747	0.610245
MYSE	0.060241	0.65	0.7	0.3	1	0.957831	0.446139
MYSO	0.138554	0.882335	0.130435	0.869565	0.895105	0.891566	0.627338
PESU	0.054217	0.885704	0.222222	0.777778	0.993631	0.981928	0.81404

_____

Table 15. Post-identification per species confidence from maximum likelihood estimator or surrogate (see program specifications for details) for known echolocation pulses for Kaleidoscope, BCID and Echoclass for simulated Georgia dataset based on Johnson et al. (2010). Note: programs may identify echolocation pulses to a species without high confidence (> 0.1-0).

_____

Kaleidoscope				BCID				Echoclass		
ID	N	p		ID	N	p		ID	N	p
EPFU	12	6.35E-05		EPFU	8	0.000001		EPFU	5	0
LABO	13	1		LABO	15	0.000001		LABO	4	0.0812
LACI	9	0.009215		LACI	8	0.000001		LACI	7	0.0618
MYGR	35	0		MYGR	30	0.000001		MYGR	26	0
PESU	51	0		PESU	56	0.000001		PESU	41	0
NO ID	1			UNKN	1			UNKN	36	
NOISE	0									

_____

Table 16. Per species echolocation pulse assignment and overall correct classification rate for Kaleidoscope for simulated Georgia dataset based on Johnson et al. (2010). Columns in bold are known echolocation pulse assignments. Rows in italics are Kaleidoscope assignment.

_____

	EPFU	**LABO**	**LACI**	**MYGR**	**PESU**	**MISC/ODD**
EPFU	5	0	0	0	0	0
LABO	0	3	0	0	0	1
LACI	0	0	7	0	0	0
MYGR	0	0	0	25	1	0
NYHU	0	1	0	0	1	0
PESU	0	1	0	0	40	0
UNKN	6	5	2	8	14	1

%
correct   95.2381

_____

Table 17. Per species echolocation pulse assignment performance for Kaleidoscope for simulated Georgia dataset based on Johnson et al (2010). Threshold is the proportional abundance of per species echolocation pulses. AUC is area under the curve. Omission is proportion of occurrences misidentified adjusted by the threshold value. Sensitivity is the proportion of true events assigned correctly (assessment of true positive). Specificity is the proportion of non-events assigned correctly (assessment of true negative). Correct is per species overall correct classification. Kappa measures the proportion of correct classification after accounting for the probability of chance agreement.

_____

	Threshold	AUC	Omission	Sensitivity	Specificity	Correct	Kappa
EPFU	0.091667	0.995413	0	1	0.990826	0.991667	0.951923
LABO	0.083333	0.986364	0	1	0.972727	0.975	0.856
LACI	0.075	1	0	1	1	1	1
MYGR	0.275	0.988506	0	1	0.977011	0.983333	0.958974
PESU	0.466667	0.955357	0.089286	0.910714	1	0.958333	0.915825

_____

Table 18. Per species echolocation pulse assignment and overall correct classification rate for BCID for simulated Georgia dataset based on Johnson et al. (2010). Columns in bold are known echolocation pulse assignments. Rows in italics are BCID assignment

_____

	EPFU	**LABO**	**LACI**	**MYGR**	**PESU**	**MISC/ODD**
EPFU	7	0	1	0	0	0
LABO	0	11	0	0	3	2
LACI	3	0	5	0	0	0
LANO	1	0	0	0	0	0
MYGR	0	0	0	30	0	0
PESU	0	0	0	4	53	0
UNKN	0	0	0	0	1	0

%
correct   89.83051

_____

Table 19. Per species echolocation pulse assignment performance for BCID for simulated Georgia dataset based on Johnson et al. (2010). Threshold is the proportional abundance of per species echolocation pulses. AUC is area under the curve. Omission is proportion of occurrences misidentified adjusted by the threshold value. Sensitivity is the proportion of true events assigned correctly (assessment of true positive). Specificity is the proportion of non-events assigned correctly (assessment of true negative). Correct is per species overall correct classification. Kappa measures the proportion of correct classification after accounting for the probability of chance agreement

_____

		Threshold	AUC	Omission	Sensitivity	Specificity	Correct	Kappa
1	EPFU	0.091667	0.813595	0.363636	0.636364	0.990826	0.958333	0.714829
2	LABO	0.091667	0.977064	0	1	0.954128	0.958333	0.792244
3	LACI	0.05	0.903509	0.166667	0.833333	0.973684	0.966667	0.69697
4	LANO	0	NA	NA	NA	0.991667	0.991667	0
5	MYGR	0.283333	0.941176	0.117647	0.882353	1	0.966667	0.914894
6	PESU	0.466667	0.941964	0.053571	0.946429	0.9375	0.941667	0.882943

_____

Table 20. Per species echolocation pulse assignment and overall correct classification rate for Echoclass for simulated Georgia dataset based on Johnson et al. (2010). Columns in bold are known echolocation pulse assignments. Rows in italics are Echoclass assignment.

_____

	EPFU	**LABO**	**LACI**	**MYGR**	**PESU**	**MISC/ODD**
EPFU	5	0	0	0	0	0
LABO	0	3	0	0	0	1
LACI	0	0	7	0	0	0
MYGR	0	0	0	25	1	0
NYHU	0	1	0	0	1	0
PESU	0	1	0	0	40	0
UNKN	6	5	2	8	14	1

%
correct   95.2381

_____

Table 21. Per species echolocation pulse assignment performance for Echoclass for simulated Georgia dataset based on Johnson et al (2010). Threshold is the proportional abundance of per species echolocation pulses. AUC is area under the curve. Omission is proportion of occurrences misidentified adjusted by the threshold value. Sensitivity is the proportion of true events assigned correctly (assessment of true positive). Specificity is the proportion of non-events assigned correctly (assessment of true negative). Correct is per species overall correct classification. Kappa measures the proportion of correct classification after accounting for the probability of chance agreement.

	Threshold	AUC	Omission	Sensitivity	Specificity	Correct	Kappa
EPFU	0.058824	1	0	1	1	1	1
LABO	0.058824	0.79375	0.4	0.6	0.9875	0.964706	0.648276
LACI	0.082353	1	0	1	1	1	1
MYGR	0.294118	0.991667	0	1	0.983333	0.988235	0.971993
NYHU	0	NA	NA	NA	0.976471	0.976471	0
PESU	0.494118	0.964563	0.047619	0.952381	0.976744	0.964706	0.929382

Table 22. Post-identification per species confidence from maximum likelihood estimator or surrogate (see program specifications for details) for known echolocation pulses for Kaleidoscope, BCID and Echoclass for simulated South Carolina dataset based on Ford et al. (2006). Note: programs may identify echolocation pulses to a species without high confidence (> 0.10-0).

_____

Kaleidoscope				**BCID**				**Echoclass**		
ID	N	p		ID	N	p		ID	N	p
EPFU	11	7.8E-06		EPFU	8	0.000001		EPFU	5	0
LABO	45	3.7E-06		LABO	79	0.000001		LABO	56	0
LACI	9	4.29E-05		LACI	8	0.000001		LACI	7	0
MYAU	NA	NA		MYAU	0	NA		MYAU	7	0
NYHU	47	0		NYHU	31	0.000001		NYHU	21	0
PESU	36	0		PESU	31	0.000001		PESU	22	0
NOID	12			UNKN	3			UNKN	45	
NOISE	3									

_____

Table 23. Per species echolocation pulse assignment and overall correct classification rate for Kaleidoscope for simulated South Carolina dataset based on Ford et al. (2006). Columns in bold are known echolocation pulse assignments. Rows in italics are Kaleidoscope assignment.

_____

	EPFU	**LABO**	**LACI**	**MYAU**	**NYHU**	**PESU**
EPFU	5	0	0	0	0	0
LABO	0	32	0	14	8	2
LACI	0	0	7	0	0	0
MYAU	0	0	0	7	0	0
NYHU	0	7	0	0	14	0
PESU	0	1	0	0	0	21
UNKN	5	22	2	4	10	2

%
correct   72.88136

_____

Table 24. Per species echolocation pulse assignment performance for Kaleidoscope for simulated South Carolina dataset based on Ford et al (2006). Threshold is the proportional abundance of per species echolocation pulses. AUC is area under the curve. Omission is proportion of occurrences misidentified adjusted by the threshold value. Sensitivity is the proportion of true events assigned correctly (assessment of true positive). Specificity is the proportion of non-events assigned correctly (assessment of true negative). Correct is per species overall correct classification. Kappa measures the proportion of correct classification after accounting for the probability of chance agreement.

_____

	Threshold	AUC	Omission	Sensitivity	Specificity	Correct	Kappa
EPFU	0.067568	0.996377	0	1	0.992754	0.993243	0.948753
LABO	0.378378	0.729425	0.410714	0.589286	0.869565	0.763514	0.47719
LACI	0.060811	1	0	1	1	1	1
NYHU	0.202703	0.698023	0.366667	0.633333	0.762712	0.736486	0.326959
PESU	0.168919	0.93122	0.04	0.96	0.902439	0.912162	0.733813

_____

Table 25. Per species echolocation pulse assignment and overall correct classification rate for BCID for simulated South Carolina dataset based on Ford et al. (2006). Columns in bold are known echolocation pulse assignments. Rows in italics are BCID assignment.

_____

	EPFU	LABO	LACI	MYAU	NYHU	PESU
EPFU	7	1	1	0	0	0
LABO	0	51	0	18	11	0
LACI	3	0	5	0	0	0
NYHU	0	5	0	4	22	0
PESU	0	5	0	2	0	24
UNKN	0	1	0	1	0	1

%
correct   68.55346

_____

Table 26. Per species echolocation pulse assignment performance for BCID for simulated South Carolina dataset based on Ford et al (2006). Threshold is the proportional abundance of per species echolocation pulses. AUC is area under the curve. Omission is proportion of occurrences misidentified adjusted by the threshold value. Sensitivity is the proportion of true events assigned correctly (assessment of true positive). Specificity is the proportion of non-events assigned correctly (assessment of true negative). Correct is per species overall correct classification. Kappa measures the proportion of correct classification after accounting for the probability of chance agreement.

_____

	Threshold	AUC	Omission	Sensitivity	Specificity	Correct	Kappa
EPFU	0.062893	0.843289	0.3	0.7	0.986577	0.968553	0.720169
LABO	0.389937	0.761806	0.177419	0.822581	0.701031	0.748428	0.497551
LACI	0.037736	0.906863	0.166667	0.833333	0.980392	0.974843	0.701408
NYHU	0.207547	0.797619	0.333333	0.666667	0.928571	0.874214	0.608856
PESU	0.150943	0.974074	0	1	0.948148	0.955975	0.846631

_____

Table 27. Per species echolocation pulse assignment and overall correct classification rate for Echoclass for simulated South Carolina dataset based on Ford et al. (2006). Columns in bold are known echolocation pulse assignments. Rows in italics are Echoclass assignment.

_____

	EPFU	**LABO**	**LACI**	**MYAU**	**NYHU**	**PESU**
EPFU	5	0	0	0	0	0
LABO	0	32	0	14	8	2
LACI	0	0	7	0	0	0
MYAU	0	0	0	7	0	0
NYHU	0	7	0	0	14	0
PESU	0	1	0	0	0	21
UNKN	5	22	2	4	10	2

%
correct   72.88136

_____

Table 28. Per species echolocation pulse assignment performance for Echoclass for simulated South Carolina dataset based on Ford et al (2006). Threshold is the proportional abundance of per species echolocation pulses. AUC is area under the curve. Omission is proportion of occurrences misidentified adjusted by the threshold value. Sensitivity is the proportion of true events assigned correctly (assessment of true positive). Specificity is the proportion of non-events assigned correctly (assessment of true negative). Correct is per species overall correct classification. Kappa measures the proportion of correct classification after accounting for the probability of chance agreement.

_____

	Threshold	AUC	Omission	Sensitivity	Specificity	Correct	Kappa
EPFU	0.042373	1	0	1	1	1	1
LABO	0.338983	0.746154	0.2	0.8	0.692308	0.728814	0.448598
LACI	0.059322	1	0	1	1	1	1
MYAU	0.177966	0.666667	0.666667	0.333333	1	0.881356	0.451163
NYHU	0.186441	0.781723	0.363636	0.636364	0.927083	0.872881	0.573494
PESU	0.194915	0.951259	0.086957	0.913043	0.989474	0.974576	0.917636

_____

Table 29.   Post-identification per species confidence from maximum likelihood estimator or surrogate (see program specifications for details) for known echolocation pulses for Kaleidoscope, BCID and Echoclass for simulated New York and West Virginia datasets based on Ford et al. (2011, 2005) with Northern long-eared bat (*Myotis septentrionalis*; MYSE) and/or Indiana bat (*Myotis sodalis*; MYSO) alternatively removed or retained. High confidence of predicted presence indicted at $p < 0.1$. False positives (gray), false negatives (yellow) noted.

			Kaleidoscope	BCID	Echoclass
New York	No MYSE	MYSE presence	1	1	1
New York	No MYSE	MYSO presence	0.004963	0.000001	0
New York	No MYSO	MYSE presence	1	0.000001	0
New York	No MYSO	MYSO presence	0.028229	0.000001	0
West Virginia	No MYSE	MYSE presence	1	0.031761	1
West Virginia	No MYSE	MYSO presence	0.0000005	0.000001	0
West Virginia	No MYSO	MYSE presence	0.31107	0.000001	0
West Virginia	No MYSO	MYSO presence	0.000786	0.000001	0