



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

Virginia Cooperative Fish and Wildlife Research Unit
106 Cheatham Hall
Virginia Tech
Blacksburg, VA 24061

Fax: (540) 231-7580
Phone: (540) 231-5927
Email: wmford@vt.edu

January 19, 2015

Memo to: Mike Armstrong, Andrew King and Robin Niver

From: Mark Ford *Mark*

Subject: Echolocation Identification Software Results – Phase II

Following our initial testing reported to U.S. Fish and Wildlife Service on September 15 and after discussions with the Service and acoustical software vendors, we agreed to retest subsequent software submissions of BCID and Kaleidoscope. Dr. Eric Britzke, developer of Echoclass, based on the satisfactory performance of his original submission and the estimated time to new modifications or versions opted to forego subsequent testing. The aim of this round of testing was to examine potential improvements on decreasing the rate of false positives and false negatives for Indiana bats MYSO (*Myotis sodalis*) and Northern long-eared bats MYSE (*Myotis septentrionalis*). Other species examined were big brown bats EPFU (*Eptesicus fuscus*), eastern red bats LABO (*Lasiurus borealis*), hoary bats LACI (*Lasiurus cinereus*), silver-haired bats LANO (*Lasionycteris noctivagans*), small-footed bats MYLE (*Myotis leibii*), little brown bat MYLU (*Myotis lucifugus*) and tri-colored bats PESU (*Periomyotis subflavus*). Following the same testing and analytical protocol as that described in the earlier 2014 round of testing, and using for this exercise the same simulated New York and West Virginia datasets, Alex Silvis and I reran all species identification, classification rates, and per species measures of omission, sensitivity and specificity, etc., for BCID 2.7b and Kaleidoscope 2.2.2. BCID was run in default settings, whereas Kaleidoscope was examined using default or neutral (0) settings as well as conservative/accurate (+1) and liberal/sensitive settings (-1). Testing of the different sensitivity settings in Kaleidoscope was done per the request of the developer and the Fish and Wildlife Service. BCID forwent providing recommended settings for testing. In addition, we also reanalyzed these programs by alternatively removing Indiana bats and northern long-eared bats as a follow-up test on classification error and the propensity for false positive values. **Overall, both programs improved and both programs continue to meet or exceed the performance standards as set forth by the Service, particularly now with regard to northern long-eared bat discrimination** (Tables 1-18). Both the conservative and liberal settings of Kaleidoscope still had relatively high omission rates for northern long-eared bats, however. But owing to the poor-quality echolocation characteristics of that species, that is not necessarily surprising. When looking at datasets with Indiana bats and northern long-eared bats removed, both programs (including all three versions of Kaleidoscope) still produced false positive errors for Indiana bats (Table 19). Both programs avoided false positives for northern long-eared bats when northern long-eared bats were removed from the data. In terms of overall accuracy relative to what echolocation pulses were identified as *a priori*, the Kaleidoscope liberal (-1) filter setting produced the most accurate results. We believe it might be worthwhile to discuss scenarios where different Kaleidoscope settings are used based on previous survey work or knowledge of Myotid distribution. Accordingly, these two new software versions should supersede their earlier submission. Efforts to continue to improve performance and refine accuracy per our earlier recommendations would still be prudent.

cc: A. Silvis

Table 1. Post-identification per species confidence (p) from maximum likelihood estimator or surrogate (see program specifications for details) and total bat pass count for known echolocation pulses for BCID 2.7b and Kaleidoscope 2.2.2 on neutral setting (0), conservative setting (+1) and liberal setting (-1) for simulated New York dataset based on Ford et al. (2011). ND is total absence of probability of detection. Note: programs may identify echolocation pulses to a species without high confidence (> 0.1-0).

BCID			Kaleidoscope 0			Kaleidoscope +1			Kaleidoscope -1		
ID	p	N	ID	p	N	ID	p	N	ID	p	N
EPFU	0.000001	21	EPFU	0.002787	14	EPFU	0.0001118	11	EPFU	0.066845	15
LANO	0.000001	10	LABO	0	26	LABO	0	26	LABO	0	26
LABO	0.000001	25	LACI	0	24	LACI	0	18	LACI	0	24
LACI	0.000001	22	LANO	0.000016	23	LANO	0.0000001	21	LANO	0.000011	23
MYLE	0.000001	8	MYLE	0.007305	6	MYLE	0.0163415	3	MYLE	0.037367	7
MYLU	0.000001	79	MYLU	0	57	MYLU	0	20	MYLU	0	67
MYSE	0.000001	12	MYSE	0.000987	9	MYSE	ND	0	MYSE	0.001060	10
MYSO	0.000001	41	MYSO	0	29	MYSO	0	11	MYSO	0	52
PESU	0.000001	13	PESU	0.000024	11	PESU	0.001146	6	PESU	0.001072	11
UNKN		4	UNKN		40	UNKN		123	UNKN		4

Table 3. Per species echolocation pulse assignment and overall correct classification rate for Kaleidoscope 2.2.2 on neutral setting (0) 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	PESU	MISC/ODD
<i>EPFU</i>	14	0	0	0	0	0	0	0	0	0
<i>LABO</i>	0	18	0	0	0	6	0	0	1	1
<i>LACI</i>	2	0	22	0	0	0	0	0	0	0
<i>LANO</i>	10	2	0	10	0	1	0	0	0	0
<i>MYLE</i>	0	0	0	0	6	0	0	0	0	0
<i>MYLU</i>	0	0	0	0	0	51	0	4	0	2
<i>MYSE</i>	0	0	0	0	2	0	7	0	0	0
<i>MYSO</i>	0	0	0	0	0	13	2	14	0	0
<i>PESU</i>	0	1	0	0	0	1	0	0	9	0
<i>UNKN</i>	0	1	0	0	2	19	8	9	0	1

% correct	77.04
-----------	-------

Table 4. Per species echolocation pulse assignment and overall correct classification rate for Kaleidoscope 2.2.2 on conservative setting (+1) 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	PESU	MISC/ODD
<i>EPFU</i>	11	0	0	0	0	0	0	0	0	0
<i>LABO</i>	0	18	0	0	0	6	0	0	1	1
<i>LACI</i>	0	0	18	0	0	0	0	0	0	0
<i>LANO</i>	10	0	0	10	0	1	0	0	0	0
<i>MYLE</i>	0	0	0	0	3	0	0	0	0	0
<i>MYLU</i>	0	0	0	0	0	18	0	0	0	2
<i>MYSE</i>	0	0	0	0	0	0	0	0	0	0
<i>MYSO</i>	0	0	0	0	0	5	0	6	0	0
<i>PESU</i>	0	0	0	0	0	0	0	0	6	0
<i>UNKN</i>	5	4	4	0	7	61	17	21	3	1

%
correct 79.65

Table 5. Per species echolocation pulse assignment and overall correct classification rate for Kaleidoscope 2.2.2 on liberal setting (-1) 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	PESU	MISC/ODD
<i>EPFU</i>	14	1	0	0	0	0	0	0	0	0
<i>LABO</i>	0	18	0	0	0	6	0	0	1	1
<i>LACI</i>	2	0	22	0	0	0	0	0	0	0
<i>LANO</i>	10	2	0	10	0	1	0	0	0	0
<i>MYLE</i>	0	0	0	0	6	0	0	1	0	0
<i>MYLU</i>	0	0	0	0	0	60	0	5	0	2
<i>MYSE</i>	0	0	0	0	2	0	8	0	0	0
<i>MYSO</i>	0	0	0	0	1	20	9	21	0	1
<i>PESU</i>	0	1	0	0	0	1	0	0	9	0
<i>UNKN</i>	0	0	0	0	1	3	0	0	0	0

%
correct 72.73

Table 6. Per species echolocation pulse assignment performance for BCID 2.7b 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). Kappa measures the proportion of correct classification after accounting for the probability of chance agreement.

	Threshold	AUC	Omission	Sensitivity	Specificity	Kappa
EPFU	0.11	0.86	0.27	0.73	0.99	0.79
LABO	0.10	0.89	0.18	0.82	0.97	0.74
LACI	0.07	0.96	0.06	0.94	0.97	0.80
LANO	0.04	0.84	0.30	0.70	0.99	0.69
MYLE	0.04	0.90	0.20	0.80	1.00	0.88
MYLU	0.39	0.88	0.19	0.81	0.96	0.78
MYSE	0.07	0.86	0.27	0.73	1.00	0.80
MYSO	0.12	0.88	0.15	0.85	0.91	0.62
PESU	0.04	0.94	0.10	0.90	0.98	0.77

Table 7. Per species echolocation pulse assignment performance for Kaleidoscope 2.2.2 on neutral setting (0) 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). Kappa measures the proportion of correct classification after accounting for the probability of chance agreement.

	Threshold	AUC	Omission	Sensitivity	Specificity	Kappa
EPFU	0.13	0.77	0.46	0.54	1.00	0.67
LABO	0.11	0.91	0.14	0.86	0.96	0.74
LACI	0.11	0.99	0.00	1.00	0.99	0.95
LANO	0.05	0.97	0.00	1.00	0.93	0.58
MYLE	0.04	0.88	0.25	0.75	1.00	0.85
MYLU	0.36	0.83	0.29	0.71	0.95	0.69
MYSE	0.05	0.88	0.22	0.78	0.99	0.77
MYSO	0.09	0.85	0.22	0.78	0.92	0.54
PESU	0.05	0.94	0.10	0.90	0.99	0.85

Table 8. Per species echolocation pulse assignment performance for Kaleidoscope 2.2.2 on conservative setting (+1) 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). Kappa measures the proportion of correct classification after accounting for the probability of chance agreement.

	Threshold	AUC	Omission	Sensitivity	Specificity	Kappa
EPFU	0.18	0.76	0.48	0.52	1.00	0.64
LABO	0.16	0.96	0.00	1.00	0.92	0.78
LACI	0.16	1.00	0.00	1.00	1.00	1.00
LANO	0.09	0.95	0.00	1.00	0.90	0.60
MYLE	0.03	1.00	0.00	1.00	1.00	1.00
MYLU	0.26	0.79	0.40	0.60	0.98	0.65
MYSE	-	-	1.00	-	-	-
MYSO	0.05	0.98	0.00	1.00	0.95	0.68
PESU	0.06	0.93	0.14	0.86	1.00	0.92

Table 9. Per species echolocation pulse assignment performance for Kaleidoscope 2.2.2 on liberal setting (-1) 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). Kappa measures the proportion of correct classification after accounting for the probability of chance agreement.

	Threshold	AUC	Omission	Sensitivity	Specificity	Kappa
EPFU	0.11	0.77	0.46	0.54	1.00	0.65
LABO	0.09	0.89	0.18	0.82	0.96	0.72
LACI	0.09	1.00	0.00	1.00	0.99	0.95
LANO	0.04	0.97	0.00	1.00	0.94	0.58
MYLE	0.04	0.83	0.33	0.67	1.00	0.74
MYLU	0.37	0.82	0.32	0.68	0.95	0.67
MYSE	0.07	0.73	0.53	0.47	0.99	0.57
MYSO	0.11	0.81	0.22	0.78	0.85	0.45
PESU	0.04	0.95	0.10	0.90	0.99	0.85

Table 10. Post-identification per species confidence from maximum likelihood estimator or surrogate (see program specifications for details) for known echolocation pulses for BCID 2.7b and Kaleidoscope 2.2.2 on neutral setting (0), conservative setting (+1) and liberal setting (-1) West Virginia dataset based on Ford et al. (2005). ND is total absence of probability of detection. Note: programs may identify echolocation pulses to a species without high confidence (> 0.1-0).

BCID			Kaleidoscope 0			Kaleidoscope +1			Kaleidoscope -1		
ID	p	N	ID	p	N	ID	p	N	ID	p	N
EPFU	0.000001	24	EPFU	0.000343	15	EPFU	0.000108	11	EPFU	0.001612	18
LANO	0.000099	6	LABO	0.000000	37	LABO	0.000000	37	LABO	0.000000	37
LABO	0.000001	42	LACI	0.000000	25	LACI	0.000000	19	LACI	0.000000	25
LACI	0.000001	21	LANO	0.004615	18	LANO	0.000187	16	LANO	0.004052	18
MYLE	0.039996	3	MYLE	1.000000	1	MYLE	ND	0	MYLE	0.917026	3
MYLU	0.000001	53	MYLU	0.000000	36	MYLU	0.000002	18	MYLU	0.000001	44
MYSE	0.000001	18	MYSE	0.000000	14	MYSE	0.215773	3	MYSE	0.000000	15
MYSO	0.000001	39	MYSO	0.000000	38	MYSO	0.000000	16	MYSO	0.000000	52
PESU	0.000001	17	PESU	0.000000	18	PESU	0.000000	16	PESU	0.000001	18
UNKN		2	UNKN		31	UNKN		97	UNKN		3

Table 15. Per species echolocation pulse assignment performance for BCID 2.7b 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). Kappa measures the proportion of correct classification after accounting for the probability of chance agreement.

	Threshold	AUC	Omission	Sensitivity	Specificity	Kappa
EPFU	0.12	0.86	0.27	0.73	0.98	0.75
LABO	0.18	0.90	0.15	0.85	0.96	0.79
LACI	0.08	0.96	0.06	0.94	0.98	0.83
LANO	0.02	0.69	0.60	0.40	0.98	0.35
MYLE	0.05	0.65	0.70	0.30	1.00	0.45
MYLU	0.27	0.91	0.17	0.83	0.98	0.85
MYSE	0.08	0.97	0.06	0.94	0.99	0.91
MYSO	0.12	0.95	0.04	0.96	0.93	0.75
PESU	0.07	0.89	0.20	0.80	0.98	0.73

Table 16. Per species echolocation pulse assignment performance for Kaleidoscope 2.2.2 on neutral setting (0) 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). Kappa measures the proportion of correct classification after accounting for the probability of chance agreement.

	Threshold	AUC	Omission	Sensitivity	Specificity	Kappa
EPFU	0.13	0.77	0.46	0.54	0.99	0.65
LABO	0.19	0.91	0.16	0.84	0.97	0.82
LACI	0.11	0.99	0.00	1.00	0.98	0.93
LANO	0.02	0.96	0.00	1.00	0.93	0.34
MYLE	0.04	0.56	0.89	0.11	1.00	0.19
MYLU	0.25	0.82	0.34	0.66	0.98	0.71
MYSE	0.07	0.86	0.27	0.73	0.98	0.74
MYSO	0.09	0.92	0.06	0.94	0.89	0.55
PESU	0.07	0.96	0.07	0.93	0.98	0.84

Table 17. Per species echolocation pulse assignment performance for Kaleidoscope 2.2.2 on conservative setting (+1) 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). Kappa measures the proportion of correct classification after accounting for the probability of chance agreement.

	Threshold	AUC	Omission	Sensitivity	Specificity	Kappa
EPFU	0.15	0.76	0.48	0.52	1.00	0.65
LABO	0.25	0.95	0.06	0.94	0.95	0.87
LACI	0.13	1.00	0.00	1.00	0.99	0.97
LANO	0.03	0.95	0.00	1.00	0.91	0.37
MYLE	-	-	1.00	-	-	-
MYLU	0.19	0.80	0.38	0.62	0.98	0.68
MYSE	0.03	0.75	0.50	0.50	0.99	0.56
MYSO	0.07	0.97	0.00	1.00	0.94	0.69
PESU	0.11	0.96	0.07	0.93	0.98	0.89

Table 18. Per species echolocation pulse assignment performance for Kaleidoscope 2.2.2 on liberal setting (-1) 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). Kappa measures the proportion of correct classification after accounting for the probability of chance agreement.

	Threshold	AUC	Omission	Sensitivity	Specificity	Kappa
EPFU	0.11	0.76	0.46	0.54	0.98	0.60
LABO	0.17	0.89	0.20	0.80	0.97	0.80
LACI	0.10	0.99	0.00	1.00	0.99	0.93
LANO	0.02	0.87	0.20	0.80	0.94	0.32
MYLE	0.04	0.55	0.89	0.11	0.99	0.15
MYLU	0.25	0.81	0.34	0.66	0.97	0.67
MYSE	0.10	0.75	0.48	0.52	0.99	0.60
MYSO	0.11	0.83	0.19	0.81	0.85	0.46
PESU	0.07	0.96	0.07	0.93	0.98	0.84

Table 19. Post-identification per species confidence from maximum likelihood estimator or surrogate for known echolocation pulses for BCID 2.7b Kaleidoscope 2.2.2 on neutral setting (0), conservative setting (+1) and liberal setting (-1) 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 indicated at $p < 0.1$. ND is total absence of probability of detection. False positives (bold), false negatives (italics) noted.

			BCID	Kaleidoscope 0	Kaleidoscope +1	Kaleidoscope -1
New York	No MYSE	MYSE presence	0.62904	1	ND	0.99568
New York	No MYSE	MYSO presence	0.00006	0	0	0
New York	No MYSO	MYSE presence	0.00006	0.00039	ND	0.0003
New York	No MYSO	MYSO presence	0.00006	0.00002	0.00382	0
West Virginia	No MYSE	MYSE presence	0.03111	0.60729	0.96070	0.61934
West Virginia	No MYSE	MYSO presence	0.00006	0	0	0
West Virginia	No MYSO	MYSE presence	0.00006	0	<i>0.21060</i>	0
West Virginia	No MYSO	MYSO presence	0.00006	0	0.00022	0