Supplement for Open-Source machine learning BANTER acoustic classification of beaked whale echolocation pulses

Shannon Rankin

2023-08-30

# NAtlantic

Results for the NAtlantic EC (only) model are provided in [Figure 1](#fig-natlantic_EC) (Detector Model: *sampsize* = 3, *ntree* = 10,000; Event Model: *sampsize* = 3, *ntree* = 10,000).

|  |
| --- |
| Figure 1: BANTER classification results from the NAtlantic dataset including echolocation pulses (EC). Confusion matrix (a) provides the percent correct classification for each species (pct.correct), lower confidence intervals (LCI\_0.95), upper confidence intervals (UCI\_0.95), and priors (expected error rate). Proximity plot (b) for species events from BANTER model (central dot color represents true species identity; color of circle surrounding dot represents BANTER species classification). Heat map (c) for ranks of ten most important variables; colors scale from most important predictors (dark red) to least important predictors (dark blue). Vote Plot (d) shows the vote distribution for each event (vertical slice) for each species; distribution of votes by species is shown by their representative color. |

# 

# SAtlantic

Results for the SAtlantic EC (only) model are provided in [Figure 2](#fig-satlantic_EC) (Detector Model: *sampsize* = 8, *ntree* = 10,000; Event Model: *sampsize* = 1, *ntree* = 100,000).

|  |
| --- |
| Figure 2: BANTER classification results from the SAtlantic dataset including echolocation pulses (EC). Confusion matrix (a) provides the percent correct classification for each species (pct.correct), lower confidence intervals (LCI\_0.95), upper confidence intervals (UCI\_0.95), and priors (expected error rate). Proximity plot (b) for species events from BANTER model (central dot color represents true species identity; color of circle surrounding dot represents BANTER species classification). Heat map (c) for ranks of ten most important variables; colors scale from most important predictors (dark red) to least important predictors (dark blue). Vote Plot (d) shows the vote distribution for each event (vertical slice) for each species; distribution of votes by species is shown by their representative color. |

Results for the SAtlantic EC\_IPI\_ALT model are provided in [Figure 3](#fig-satlantic_EC_IPI_ALT) (Detector Model: *sampsize* = 4, *ntree* = 10,000; Event Model: *sampsize* = 3, *ntree* = 100,000).

|  |
| --- |
| Figure 3: BANTER classification results from the alternative SAtlantic dataset including echolocation pulses and inter-pulse interval with a large sampsize (EC\_IPI\_ALT). Confusion matrix (a) provides the percent correct classification for each species (pct.correct), lower confidence intervals (LCI\_0.95), upper confidence intervals (UCI\_0.95), and priors (expected error rate). Proximity plot (b) for species events from BANTER model (central dot color represents true species identity; color of circle surrounding dot represents BANTER species classification). Heat map (c) for ranks of ten most important variables; colors scale from most important predictors (dark red) to least important predictors (dark blue). Vote Plot (d) shows the vote distribution for each event (vertical slice) for each species; distribution of votes by species is shown by their representative color. |

# 

# Hawaii

Results for the Hawaii EC\_IPI model are provided in [Figure 4](#fig-hawaii_EC_IPI) (Detector Model: *sampsize* = 5, *ntree* = 100,000; Event Model: *sampsize* = 4, *ntree* = 10,000).

|  |
| --- |
| Figure 4: BANTER classification results from the Hawaii dataset including echolocation pulses and inter-pulse interval (EC\_IPI). Confusion matrix (a) provides the percent correct classification for each species (pct.correct), lower confidence intervals (LCI\_0.95), upper confidence intervals (UCI\_0.95), and priors (expected error rate). Proximity plot (b) for species events from BANTER model (central dot color represents true species identity; color of circle surrounding dot represents BANTER species classification). Heat map (c) for ranks of ten most important variables; colors scale from most important predictors (dark red) to least important predictors (dark blue). Vote Plot (d) shows the vote distribution for each event (vertical slice) for each species; distribution of votes by species is shown by their representative color. |

# 

# EPacific

Results for the EPacific EC (only) model are provided in [Figure 5](#fig-epacific_EC) (Detector Model: *sampsize* = 4, *ntree* = 10,000; Event Model: *sampsize* = 4, *ntree* = 10,000).

|  |
| --- |
| Figure 5: BANTER classification results from the EPacific dataset including echolocation pulses (EC). Confusion matrix (a) provides the percent correct classification for each species (pct.correct), lower confidence intervals (LCI\_0.95), upper confidence intervals (UCI\_0.95), and priors (expected error rate). Proximity plot (b) for species events from BANTER model (central dot color represents true species identity; color of circle surrounding dot represents BANTER species classification). Heat map (c) for ranks of ten most important variables; colors scale from most important predictors (dark red) to least important predictors (dark blue). Vote Plot (d) shows the vote distribution for each event (vertical slice) for each species; distribution of votes by species is shown by their representative color. |

Results for the EPacific EC (only) model are provided in [Figure 6](#fig-epacific_EC_IPI) (Detector Model: *sampsize* = 3, *ntree* = 10,000; Event Model: *sampsize* = 4, *ntree* = 10,000).

|  |
| --- |
| Figure 6: BANTER classification results from the EPacific dataset including echolocation pulses and inter-pulse interval (EC\_IPI). Confusion matrix (a) provides the percent correct classification for each species (pct.correct), lower confidence intervals (LCI\_0.95), upper confidence intervals (UCI\_0.95), and priors (expected error rate). Proximity plot (b) for species events from BANTER model (central dot color represents true species identity; color of circle surrounding dot represents BANTER species classification). Heat map (c) for ranks of ten most important variables; colors scale from most important predictors (dark red) to least important predictors (dark blue). Vote Plot (d) shows the vote distribution for each event (vertical slice) for each species; distribution of votes by species is shown by their representative color. |