

Kaleidoscope Pro 5 Cluster Analysis For Bats

Welcome to Kaleidoscope Pro from Wildlife Acoustics.

This video explores the clustering features of Kaleidoscope Pro as used for detection of bats.

Kaleidoscope Pro has an Auto-ID function for bats, but there are times one might not want to rely on Auto-ID. Perhaps what I'd like to do is examine bat recordings to group different species together for further analysis. Kaleidoscope Pro can do a Cluster Analysis batch process to provide the solution. Let's take a look.

The source recordings for this example were made over the course of a year. This is a great deal of data. Let's find out what type of bats showed up in these recordings in 2016.

Under the Cloud tab in Kaleidoscope Pro I've directed the batch process Input folder to access a folder in a Wildlife Acoustics Managed Cloud Account. These recordings were made by a colleague, and I have access to his files in the Cloud account for analysis. For the output results I've selected a folder in the Managed Cloud Account that I can use for output results.

I've made sure that Kaleidoscope Pro is in Bat Analysis mode. Under the Signal Parameters tab I'm using the default settings which include a Maximum inter-syllable gap of 500ms. This sets the maximum distance between detected signals before a detected signal is considered to be complete. If I lower this setting I can isolate individual bat calls. A larger inter-syllable gap means Kaleidoscope Pro will look at sequences of calls. For this analysis I want Kaleidoscope Pro to analyze sequences of calls that have a minimum of two pulses that are no further than 500ms apart, so I'll go with the larger default inter-syllable gap.

The first step is to do a basic analysis of the raw data. Under the Cluster Analysis tab I've chosen the option to Scan and cluster recordings to create cluster.kcs and cluster.csv files. I'm also using the default settings under this tab.

I'll go ahead and press the Process Files button. Because both input and output folders are in the Cloud, I have the option to run the cluster analysis batch process using either Cloud-based Computing, or running the batch process on my local computer. For this example, I'll select the option to use Cloud-based computing. I've entered my email address for notifications. I'll receive an automatic notification when the batch process is started and a second notification when the batch is completed.

Kaleidoscope Pro uploads the Cluster Analysis parameters to the Managed Cloud Account. A version of Kaleidoscope Pro is built into the Wildlife Acoustics Managed Cloud Account, and will run the batch processing on-line.

Video Script



Once the Cloud-based batch process is complete I'll direct Kaleidoscope Pro to the designated output folder in the Managed Cloud Account where I see the newly created cluster.csv output file. When I double click on this file Kaleidoscope Pro opens the Viewer and Results window to display the Cluster Analysis results.

When I scroll down in the Results window I see that Kaleidoscope Pro has detected over 50,000 bat passes from the input recordings. Remember, this is a year's worth of recordings. Kaleidoscope Pro has created a total of 75 different clusters based on the input files and software settings. I'm fairly sure there are not 75 different species of bats in these recordings. Kaleidoscope Pro creates different clusters based on the statistical similarities of the recordings. Bats produce a large variety of call sequences as they adapt to their surroundings, navigate around objects such as trees, and hunt for insects. There could be combinations of multiple bats or different species flying at the same time. So these different clusters may ultimately represent the same species of bats but who happen to be doing very different things in each recording.

I'm confident I will find clusters that represent clean examples of individual species.

I've gone ahead and examined the different clusters and I've found that cluster #0, which represents the cluster with the largest number of similar detections, most likely represents Big Brown bats. I can bulk-label a cluster with a manual ID so all the vocalizations within that cluster are manually identified. That's what I'll do here.

Clusters 60 and 61 appear to be Hoary bats so I'll label those clusters accordingly. Hoary bats are migratory so I'm interested to see more information about when they were recorded.

Clusters 73 and 74 appear to be pretty good representations of Eastern Red bats, so I'll label those. The other clusters appear to contain recordings of these species, especially the common Brown Bat, but again, there could be ambiguous representations of species or recordings that contain multiple bats and other environmental noises. By labelling these clusters, we are using the clusters as a proxy for bat activity of a particular species. For Big Brown bats, while there are many clusters containing examples of Big Brown bats doing different things, we'll just label cluster.0 as a representative cluster. If I was interested in relative call activity levels, I might explore and label additional clusters as Big Brown bats as well to include those results in my analysis. But for this example, we'll just label the largest cluster.

I'll save the Results window which updates the underlying cluster.csv file with my manual species labels.

It's also possible to use clustering technology to build a simple or advanced classifier to detect patterns similar to patterns discovered in the data set. The simple classifier and advanced classifier functions may be more useful for analysis of birds and other land animals. If you'd like to learn more about simple and advanced classifiers we invite you to check out the additional tutorial movies on Cluster Analysis found at the Wildlife Acoustics web site.

Video Script



For the general analysis of what bat activity has been found in my initial input recordings the work has been done. I'll download the cluster.csv file to my desktop so I can open it in a spreadsheet application. I've created a pivot table that shows which species have been identified according to the manually assigned labels. The pivot table can show call activity by species detected on each night by date. I see lots of Big Brown bats which is to be expected. I see a smaller number of Eastern Red bats which is also to be expected.

I've narrowed the date range down to July and I see that on July 29th there is a big spike of Hoary bats. It looks like July 29th may be the night these bats migrated through the recording area. So with just a little analytical work on my part, plus the Cluster Analysis functions of Kaleidoscope Pro, I've very quickly found a major clue about when Hoary bats have flown past my recorder.

Cluster Analysis for bats is an easy to use and powerful tool included with Kaleidoscope Pro from Wildlife Acoustics.

Thank you for watching.