The AviaNZ Bioacoustic Analysis Software (v 2.0)

The AviaNZ Team

(stephen.marsland@vuw.ac.nz)

November 2019

This is the user manual for version 2.0 of the AviaNZ program. The two biggest changes in this version are that we use both mouse buttons: one to draw segments, and the other to select them (the choice as to which does what is an option, but defaults to left button for making segments and right for selecting them), and we have significantly improved the way that species recognisers are trained and stored. There are a few other improvements, but these are the biggest change. As well as this manual, there are some videos and an FAQ available at http://www.avianz.net.

We provide some introductory material about sounds and spectrograms on our web page if you want to know more about them. We also give specifications of the filter storage format and the bird lists for those that wish to change them.

We really want feedback on AviaNZ, particularly what works and what doesn't, how you would like to see it improved, and what other functionality it needs. We are more than happy to talk about our plans as well.

Contents

1   Getting Started

AviaNZ has three main modes of interaction, which are presented as options on the start-up screen:


- The first option ('Manual Processing') is described more in Section 2. It enables you to look at, listen to, and manually annotate individual audio files, as well as to train your own recognisers.

- The second option ('Batch Processing') takes whole directories (and subdirectories) of audio files and automatically segments the bird calls of selected species you have recognisers for; see Section 3.

- You can view the output of the automatic segmentation using the the third option ('Review Batch Results', see Section 4), but if you want more context on them you could also use the first ('Manual Processing') option.

- The software also produces an Excel file showing the results; these are described in more detail in Section 5.

2   Manual Processing

If you select 'Manual Processing', you will see a dialog box asking you to select a file to view. Use this in the normal way to select a sound file from a directory. The program will ask you to give names for the operator and reviewer. These are useful to keep track of who has looked at different files. Following that, you should see a screen like the one on the next page. This is the main interface for manually labelling birdcalls, training recognisers, or testing things. It can also be used for reviewing the results of automatic processing; checking for misclassifications is usually easier in 'Review Batch Results', but the manual processing mode also allows you to see if there are any calls that the recogniser missed.

The program automatically loads the first 5 minutes of a file (if it is in wav format; mp3 files lose a lot of information and should not be used, other formats should be converted to wav before use). AviaNZ assumes that sound files are recording in mono (single microphone) sound. If there are other channels, it only loads the first one. The name of the current file is shown in the title of the window (1).
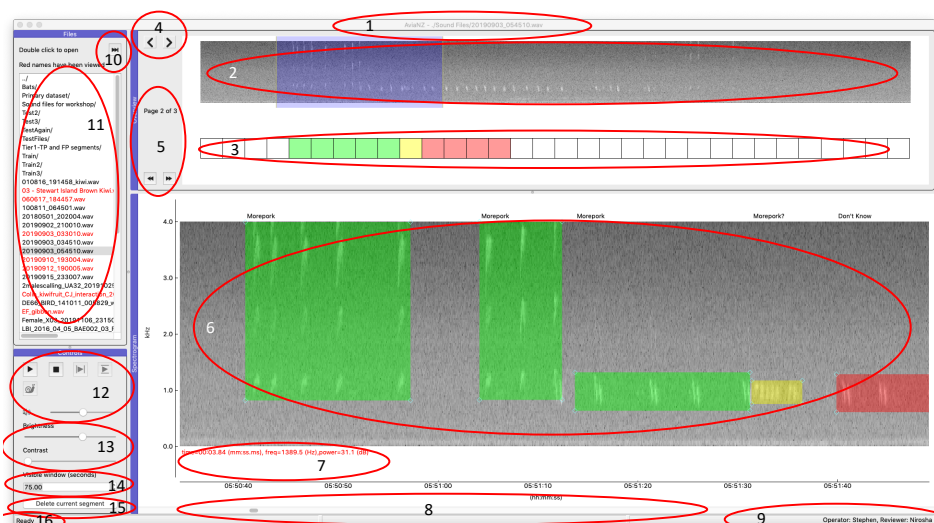
There are four separate areas of the screen within the 'Manual Processing' part of AviaNZ. Each has its name in a blue bar. They are:

Overview This shows you a picture of a 5 minute segment of the file (labelled 2 in the figure). The part you are looking at in the main plots is shown in blue. Below it there is a coloured bar (in yellow and labelled 3 in the figure). There are left and right arrow buttons (labelled 4) and double arrow buttons (labelled 5) on the left of this area. The single arrow buttons move the view in the main area along, while the double arrow buttons move to the previous 5 minutes or to the next 5 minutes of the file (if they exist).

Files (11) This is a list of files in your current directory. You can double-click on one to select it and open it. Files in red have been annotated in the program, while files in black have not. The double arrow button on the top right of this area (labelled 10) moves on to the next file.

Spectrogram (6) This is the main representation of the sound file, and the way that you add and modify annotations. It is possible to display information about where the mouse is pointing on the spectrogram (time, frequency, energy value; (7 in the figure)) by choosing 'Show pointer details in spectrogram' from the Appearance menu, and to switch it off the same way. There is also a scroll bar to move through the current page (8).

Controls These play the sound file (12), modify the appearance of the plots (13 in the figure, see Section 2.7) and enable you to change the size of the visible part of the file (14), and delete the current segment (15). For more details, see Section 2.5.

There is an optional fifth area, which is the Amplitude plot. If you want to see the amplitude plot as well, select 'Show amplitude plot' in the Appearance menu. Select the option again to hide it. You can also hide the list of files (11) and the annotation overview (2–5) in the same way. AviaNZ will remember your choice in future uses.

There are two other parts to the interface:

Menu The menu at the top of the screen.

Bar at the bottom At the very bottom of the screen there is an area (16) that gives any status updates from the program, and on the right, information about the currently selected Reviewer and Operator (9).

- You can drag the five screen areas around and reorder them if you wish, by dragging the blue bar on the top or left of them. You can also make them into their own windows by double-clicking on the blue bar. If you decide that you made a mistake doing that, then there is an option in the Actions menu to 'Put docks back' that returns them to the original configuration.

- To load a new file, either choose 'Open sound file' in the File menu, or just double click one in the Files area (double clicking on a folder will open that folder), or click on the button labelled 10 to move to the next file.

- If you want to move to a new directory, either use the '..' option at the top of the list of files to navigate around your computer's file system, or use the 'Open sound file' in the File menu.

- To restart the program, for example so that you can start doing some batch processing, choose 'Restart Program' from the File menu. To quit completely, choose 'Quit'.

2.1 Spectrogram

- The main spectrogram plot (6) shows a section of a sound file. The part you are looking at is highlighted in blue in the top (Overview: 2) picture.

- The axes of the spectrogram are time on the horizontal axis and frequencies on the vertical axis. The times will be true times if this information is available (such as when using DOC recorders), or time from start of file otherwise.

- Sometimes spectrograms don't look good initially, for example because of high noise. You can modify the brightness and contrast of the spectrogram using the two sliders (13).

- You can also use a different colour scheme, and invert the colour map (swap black and white), by choosing the relevant options in the Appearance menu.

2.2 Zooming and Scrolling

- The part of the file you can see can be changed by:
  - dragging the scroll bar below the spectrogram (8).
  - clicking on the left or right arrows on the right of the Overview picture (4).
  - dragging the blue highlight in the Overview picture itself (2).
  - clicking on any of the boxes in the bar below the Overview picture (3).
  - pressing the left or right arrow keys.

- The amount of the file that you can see (the visible width) can be changed by either:
  - dragging the ends of the blue highlight in (2).
  - changing the 'Visible window width (seconds)' in the Controls dock (14) by clicking on the up and down arrows, or typing in a new number.

- You can also view a restricted amount of the spectrogram by reducing the visible frequency band, see Section 2.7.

2.3 Moving Through Long Files

- If files are longer than 5 minutes, use the double arrows labelled (5) to move to the next or previous page, or press Shift + left or right arrow keys.

- The program tell you which page you are currently on, and many there are in total.

- There is (by default) a 10 second overlap between the pages. You can change the page size and the amount of page overlap by choosing the relevant options in the 'Interface Settings' in the Interface menu, see Section 2.9.

- The times on the axis below the spectrogram show locations in the full file.

- Note that operations like denoising and segmentation (described in Section 2.6) apply to the visible 5 minute portion of the file, not the whole file.

2.4 Manual Labelling

- To create segments, click and drag with the left mouse button on the Spectrogram.

- To select segments, click on them with the right mouse button (pressing control on the keyboard when clicking on a Mac). The segment will turn blue when it is selected.

- You can change which mouse button performs which action using the 'Mouse settings' of the 'Interface settings' in the Appearance menu.

- If you find that the colours make it hard to see the data underneath the boxes you can make them transparent using the 'Make dragged boxes transparent' option under 'Annotation' in the 'Interface settings'. You can also choose different colours if you have particular preferences.

- Segments are saved automatically, so that you can't lose your work.

Creating and Labelling Segments

- There are three options for how to create a segment. You select one of them in the 'Mouse settings' of the 'Interface settings' in the Appearance menu:

    - *(default)* Drag a limited frequency band box (i.e., click and hold the mouse button and drag the mouse to the correct end point in both time and frequency).
    - Start and stop a limited frequency band by clicking (i.e., click once at the correct time and frequency for the start of a segment, and then again at the time and frequency for the end).
    - Start and stop a full frequency band by clicking (i.e., click once at the correct time for the start of a segment, move the mouse to the end, click again, the box covers all frequencies).

- When you create a new segment, a drop-down menu will appear asking you to choose a label for that segment.

    - To choose the type of bird, click on the name.
    - If the species isn't in the list, move to 'Other'.
    - If it isn't in the second list, at the bottom of that list is a selectable box (it probably says 'Albatross'). Clicking on it provides the complete list of birds.
    - If there is something missing from there, choose Other' from it, which is at the very end. It will ask you to enter a name, as Genus (Species); e.g., 'Kiwi (Little Spotted)'.
    - If there is only a single example of the genus, you can miss out the species, e.g., 'Kakapo'.
    - New name you add will then appear in the list.
    - If you click anywhere on the screen except on a bird name in the menu then the menu will disappear and the bird type will be labelled as 'Don't Know'.
    - The bird list that we are using is based on the one that DOC use, and is meant to cover all the bird species that we know about in New Zealand.
    - You can add new species individually using 'Other' in the menus.
    - It is possible to use other bird lists using the 'Interface settings'.

- By default the lists of bird names update dynamically, so that bird types you have chosen appear at the top of the list. If you don't like that, then you can disable it in the 'Interface settings'.

- When creating a segment, you can give it the same label as the previous segment by pressing the Shift key on the keyboard when you click to finish the segment. The program will then give the segment the same label as the previous box that you labelled, without showing you the list. This is very useful when there is one bird calling repeatedly.

- You can also show that you are uncertain by pressing the Ctrl button when you click to segment (command button on a Mac computer). The names of the birds will then have a question mark after them.

5

- By default, the software only allows one bird type to be specified for each segment. However, there may be times when you wish to label multiple species in one box (for example, for the dawn chorus). In that case, choose 'Default to multiple species' in the 'Interface settings'. Now, when you choose a bird from the list it will be ticked, and the menu will not close automatically, allowing you to make multiple selections. To unselect something, click on it again. When you make a new segment, 'Don't Know' is selected by default. Choosing any other option deselects 'Don't Know'. You can reselect it if you do want that label as well.

Updating Segments

- If you select a segment that already exists by clicking on it then it will turn blue. Click on it again and the menu will reappear so that you can correct mistakes.

- Segments have blue diamonds at the corners, so that you can resize them, or you can move the whole segment by dragging it (even when it isn't selected).

- Segments can be deleted by selecting them (so that they turn blue) and then clicking on the 'Delete Current Segment' in the Controls. You can also just press the delete or backspace key on the keyboard.

- To delete all of the segments from the current file, use the 'Delete all segments' option in the Actions menu.

- You can disable the making and updating of segments (to avoid making segments by mistake) by selecting 'Make read only' from the 'Appearance' menu.

Colour Codes

- The segments that are drawn on the screen have different colours. These colours are changeable in the 'Interface settings', but by default are:

  Blue This segment is currently selected. You can play it by pressing the buttons in the Controls area (see Section 2.5), or give it a new label by clicking on it again, or delete it.

  Green This segment has been labelled with a bird name.

 **Yellow** This segment has been labelled with a bird name with a question mark.

  Red This segment has been labelled as 'Don't Know'.

- These colours match the rectangles underneath the Overview:
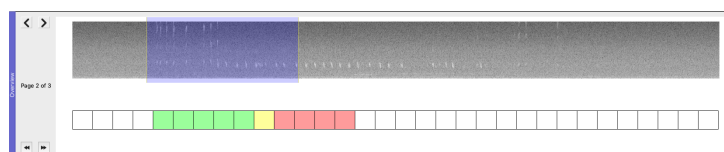
- For each 10 second segment of the file, these boxes are:

  White if there are no segments,

  Red if there are 'Don't know' segments,

  Yellow if there are question-marked segments, or

  Green if all the segments in that section are labelled with definite species.

- You can click on those rectangles, and they will update the main spectrogram plot to show that section of the file. This is a good way to move through the file quickly.

2.5 Controls
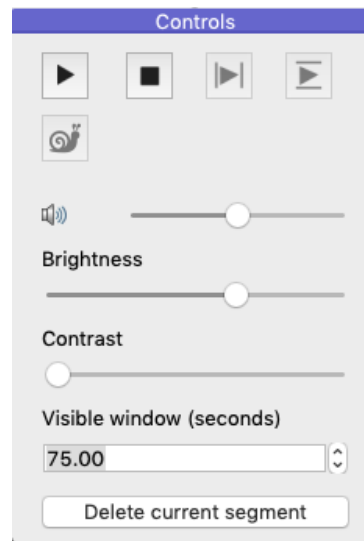
- The buttons at the top of the controls block allow you to play the sounds.

  - The top-left button is a normal play button. The button turns into a 'Pause' button while playing.
  - While the sound is playing, the light blue bar in the spectrogram plot shows where the playback is up to.
  - When the sound is paused, you can drag this bar if you want to hear a particular part of the file. Move the mouse over it, and the bar will go red. Then click and drag (using the mouse button that does not make segments, by default the right button) to move it.
  - To stop playback and have the slider return to the start of the visible section, press the Stop button.
  - When a segment is selected (so that it is blue) you can use the two play buttons to the right of the stop button to play just that small segment.
  - The difference is that the one on the left plays all the frequencies in the sound file, while the one on the right will play only the frequencies highlighted, so that you can isolate particular frequencies of a call. This is particularly helpful when there is high level of background noise in a particular frequency range, such as cicadas, or where there are overlapping calls in different frequency bands.
  - The button below them (with a picture of a snail on it) lets you play back the selected segment at different speeds. You can change the speed by clicking and holding on the button. Note that this changes the pitch of the sound.
  - You can change the volume of playback using the slider below the buttons.

- The brightness and contrast sliders change the appearance of the spectrogram, helping to see more sounds.

- The size of the visible window controls how much of the full spectrogram appears in the main window.

- The 'Delete current segment' button removes any segment that is selected (blue colour).

2.6 Menu Options

Most of the options for AviaNZ are found through the menus. There are keyboard shortcuts for many of the menu items, which can be seen in the menus themselves.

File Menu

Open sound file Produces a file dialog so that you can choose a new sound file to open.

Set operator/reviewer (Current File) Enables you to change the operator or reviewer that were specified when you started AviaNZ. The change only applies to this file.

Restart program Takes you back to the start screen so that you can access the other functions.

Quit Does what it says.

Appearance Menu

Changing appearance The first four options in the menu hide or reveal the amplitude plot, list of files, overview, and information about where the mouse is pointing in the spectrogram. The ones that are selected are marked with a tick.

Choose colour map This enables the user to select a colour map they prefer to the standard grey one.

Invert colour map By default, areas of high energy (frequencies where there is a call or other sound) are shown as the lightest colour, and low energy as dark. This can be swapped over with this option. Note that you will need to change the brightness and contrast using the controls area after inverting the colour map..

Change spectrogram parameters A set of options to modify the spectrogram. See Section 2.7 for more details. The two sliders at the bottom of the dialog enable the user to non-destructively show a limited frequency band in the spectrogram. The axis in the plot shows the range that is visible.

Mark on spectrogram There are some features that will help you spot bird calls or recognise them. We currently offer three options here: the fundamental frequency, spectral derivatives, and points of high energy. To select one, choose it from the menu; select it again to remove it. They can help to find particular types of call.

Make read only When reviewing a segmentation it is possible to click on the plots by mistake, adding further segments. This option avoids this problem. It can be turned off by selecting it again. When read only mode is on, the message section at the bottom of the screen says so.

Interface settings This option produces a dialog that enables the user to customise several things about AviaNZ. See Section 2.9 for more details.

Put docks back Returns the various screen components to their original layout if they have been moved around.

Actions Menu

These options will mostly provide dialog boxes that ask you to make choices.

Delete all segments Does what it says.

Denoise Runs some programs that try to get rid of the noise in the sound file so that the birdcalls are easier to see. See Section 2.8 for more information.

Add metadata about noise Allows the user to specify if the sound file is particularly corrupted by noise, and also to identify the type(s) if known. This can be an optional data field, or made compulsory for each file by choosing the appropriate option in the 'Interface settings'.

Segment You can ask the computer to segment the calls automatically. We currently provide 3 options. The first ('Wavelets') applies a pre-trained recogniser for a particular species, in the same way that the Batch Processing mode (Section 3) does, but just to the current page. The other two ('FIR' and 'Median Clipping') will create a segment for any significant noise in the file.

Export segments to Excel This enables the user to save a summary of the annotation of the currently opened file into an Excel workbook (in the same location as the current sound file), as is described in Section 5.

Human Review These two options provide two different ways to view the segments and their labels on the current page for easier checking. For more, see Section 4.

    All segments Show each segment regardless of label.

    Choose species Show all segments for one species.

Export current view as image Saves the spectrogram currently visible on the screen as an image, including any segments marked.

Save selected sound Saves a selected segment as a short sound file.

Recognisers Menu

Train an automated recogniser A significant new feature of this version of AviaNZ is that you can now train your own species recognisers. The process is relatively simple, and is described in Section 6.

Test a recogniser Enables you to test a recogniser by choosing a folder to run it on. See Section 6.6.

Manage recognisers Lets you rename, export, or import new recogniser for different species.
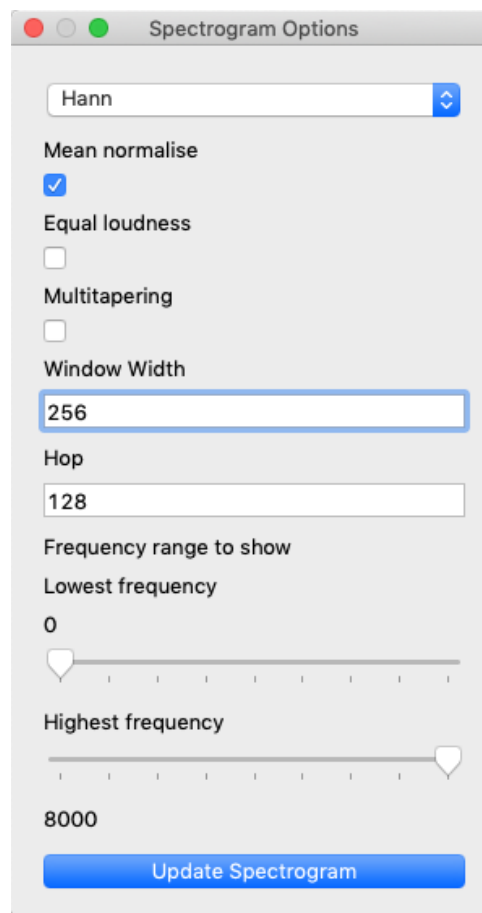
Help Menu

Help Gives access to this manual online.

Cheat sheet Links to our webpage to see examples of New Zealand bird spectrograms and calls.

About Shows you basic information about AviaNZ.

2.7 Changing the Spectrogram Computation

- Selecting 'Change spectrogram parameters' from the Appearance menu will produce the following dialog box:

- You can change some of the parameters that are used to produce the spectrogram:

  - The window function (default: Hann) controls how the spectrogram combines sounds across the time range.
  - Mean normalisation and equal loudness try to make the spectrogram energies more even.
  - Multitapering uses multiple windows to make a better estimate of the spectrogram, but performs a lot of computation, and can therefore be slow.
  - The window width and hop size control how much of the sound file makes up one spectrogram bin, and how they overlap. A big window size will improve frequency resolution, but reduce time resolution, and vice versa. The hop size controls how much the spectrogram time bins overlap.
  - The frequency range sliders let you change the visible frequency range in the main spectrogram window. This can be useful if you want to focus on only part of the range.

- If you want to know more about these options, look on the AviaNZ webpage, in the 'Technical Details' section.

- You can also just try changing them and see if it makes your spectrograms clearer.

2.8 Denoising

- When sound files are particularly noisy, it can be helpful to remove some of that noise. AviaNZ currently provides three ways to do this (via the 'Denoise' option in the Actions menu):

**Wavelets** This is our main method, and tries to preserve the bird call perfectly.

Bandpass Suppresses all frequencies outside a restricted range (which you specify) so that you can concentrate on the frequency range where calls that you are interested in can be seen.

Butterworth bandpass Another way to compute a restricted frequency range.

- You can save the denoised sound files, and undo the denoising if it does not help.

2.9 Interface Settings

- There are quite a few user-selectable options in AviaNZ, which you can choose using the 'interface settings' menu option in the Appearance menu, which will produce the following dialog box:
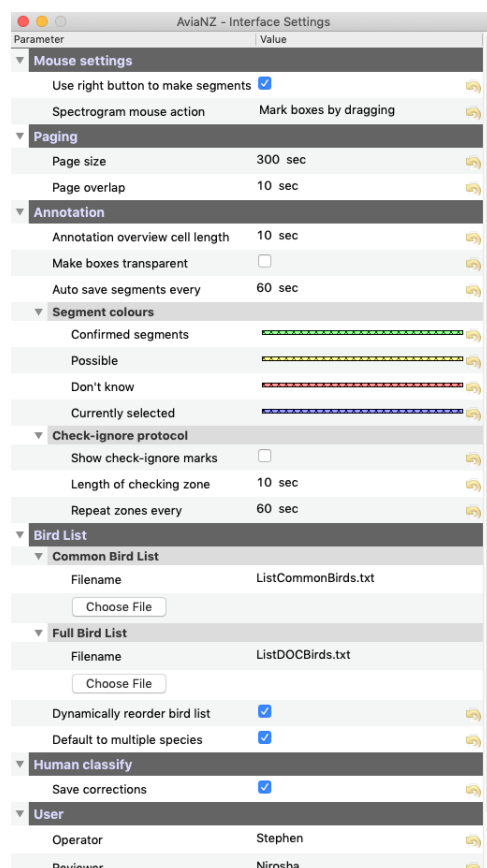
- These options include:

Mouse settings Swap which mouse button selects segments and which creates them, and change the method of creating segments (clicking or dragging).

Paging The page is the length of spectrogram loaded into AviaNZ at one time. By default it is 300 seconds (5 minutes). Note that longer times will require more memory and processing time. The amount of overlap between the pages can also be specified. The aim of the overlap is to make sure that calls aren't missed, and that segments are labelled accurately, at the page limits.

Annotation – The first option in this section changes the length of the boxes in the Overview.
- The next enables you to make the labelling boxes transparent (so that only the outline of the box is shown) if you find it hard to see what is in a segment.
- By default, AviaNZ save the segments you have made every 60 seconds. This can be changed, for example if you want to do it more often for safety.
- You can also change the colours of segments.
- AviaNZ has a check-ignore protocol for people who only annotate subsets of recordings. This puts a mark on the spectrogram in places where the user should be annotating the spectrogram, which you can control with these options.

Bird list There are two bird lists in AviaNZ: the short list of common birds, and a longer list of all possibilities. You can change these files for others (for example, to include non-New Zealand species).

You can also change the default behaviours of dynamically updating the list of common birds, and enable multiple species to be selected for a single segment. This can be useful for labelling things like the dawn chorus, but these segments will not provide good training data for new recognisers.

Human classify By default, AviaNZ saves corrections that are made using the review options (see Section 4).

User Enables you to set the operator and reviewer; this facility can also be found in the File menu. You can also change whether or not the AviaNZ window starts at full screen size, and whether or not the noise data must be completed for all files.

There are two other ways to interact with AviaNZ that are selectable via the start screen, and that are described next.

3 Batch Processing

Batch processing is for use when you have large numbers of recordings that need to be processed, for example when you collect recorders from the field. Download all of the data from the SD cards into a folder on your computer, and then start AviaNZ.
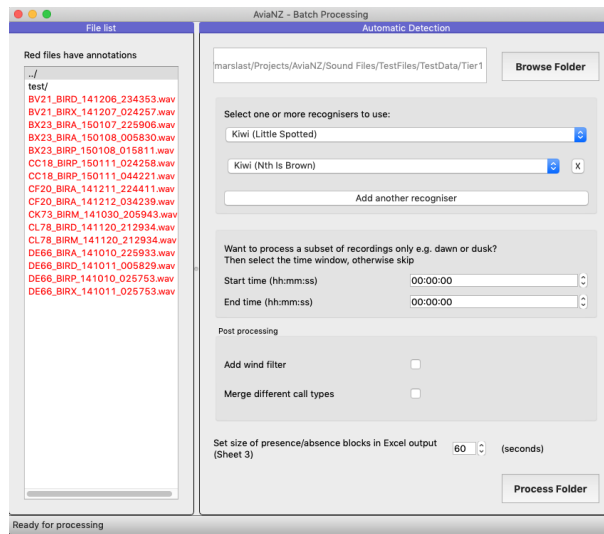

- Select 'Batch Processing' from the starting window. You will see a screen like the one above.

- Navigate to a folder containing recordings to process.

- If this folder contains subfolders, the program will work through all of the folders inside the original one.

- Note that it deletes any previous annotations in your files.

- There are two outputs from batch processing: the data files that AviaNZ uses to label segments, and Excel files annotating the presence or absence of different species (see Section 5).

- You can choose one or more species-specific recognisers from the drop-down list to apply to the sound files in order to automatically label them.

- Alternatively, you can select 'Any sounds'; in this case AviaNZ will show any sounds in the files, and label them as 'Don't Know'.

- For recordings from DOC recorders you can also choose to limit the times of the recordings.

- You can choose whether or not to try and filter out the sound of wind, and to merge call types. The first can be useful if your data was collected in a windy place, but may miss some calls otherwise.

- AviaNZ also produces an Excel spreadsheet, as is described in Section 5. One sheet of this spreadsheet shows Presence/Absence in time blocks. You can set the length of this time block here.

- Note that the 'Any sounds' option will delete the Excel files for individual species to ensure that everything is consistent. The information is then held in one spreadsheet.

- Press the 'Process Folder' button to start the program. This is a very computationally intensive process, and will take a long time (hours) if there are lots of files to process, and make your computer hard to use for anything else. It is generally best to leave it overnight.

- If you stop the processing partway through, AviaNZ will try and restart from the place it got up to last time if you restart it.

- Once it has finished, the window will give you the option to see the AviaNZ start screen again so that you can review the outputs. You can either do this in the 'Manual Processing' interface, or use the 'Review Batch Results' option, which is described next.

4   Review Batch Results

After batch processing it is important to verify the results, since AviaNZ will have created false positives (labelled calls are being your species when they are not). We provide two interfaces for checking and correcting the results.

If you select 'Review Batch Results' on the start screen then a similar menu to the 'Batch Processing' option will prompt you to select a folder with previously processed files. There are two forms of review in AviaNZ:

- If you choose 'All species' then you will see a screen like this:
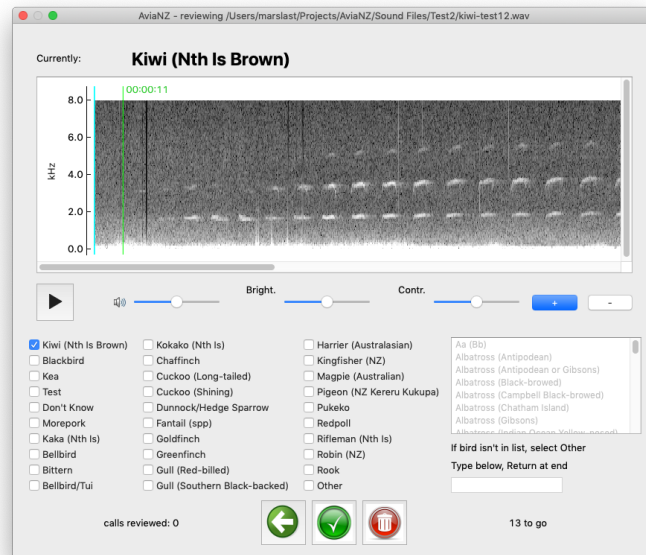
- The green bars on the spectrogram show the start and end of the call; the spectrogram includes a couple of seconds of context sound too.
- Make the window larger if the spectrogram is too large for the window, or use the '+' and '-' buttons to zoom.
- You can play the sound with the play button.
- You can also change the spectrogram brightness and contrast to make it easier to review.
- If the label is correct, click on the green tick, which will make the next image load.
- If the label is wrong, select the correct label before clicking the tick button.
- If you have allowed multiple species selection (in the 'Interface setting' in the Manual Interface) you can pick several options.
- To delete a segment, click on the red dustbin button.
- To move back to the previous one, click the back arrow; note that this will not save your current changes.

- If you select a particular species, you will be asked to choose a species from those found in the current folder.

- You will then be shown a screen like this:
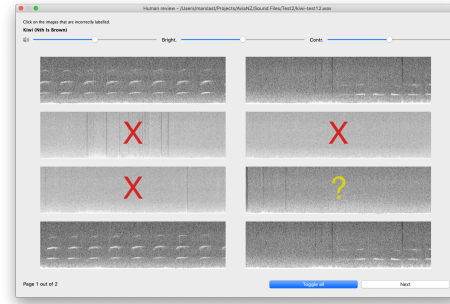
- This is a set of spectrograms from each file that have been labelled as that species.
- Make the window larger to see more of them, (any that do not fit will be shown on another page).
- For each segment that is wrongly labelled, click on its picture (it will be marked with a cross).
- If you click on it again, the cross will become a question mark to show that you are unsure, and if you click again, it will return to having no mark.
- Segments marked with a cross will be deleted, those with a question mark will be identified as unsure by having a question mark added to those label, while those left plain will be confirmed as correct.
- You can change the brightness and contrast here as well.
- You can also play the sounds by hovering on the image, and then clicking on the play button at the top-left of the image when it appears.
- The 'Toggle all' button cycles all of the spectrograms through the cross-question mark-OK stage, so that you don't have to click on every button if there are a lot of errors.
- Click on 'Next' to move on to the next screen, which will either be more spectrograms from this file, or move on to the next file.

- You can also see these two types of review from the 'Manual Processing' interface by selecting 'Human review' and then either 'All segments' or 'Choose species'.

5 Outputs

The AviaNZ program aims to provide detailed and easy-to-review outputs. An Excel file (with the name 'DetectionSummary' followed by species selected) will be generated with three sheets of outputs in the same directory as the files. If this file already exists, the new results will be appended to the end of each sheet.

The three sheets of the Excel workbook are:

1. start and end times of each birdcall detected

2. presence/absence of the target species (or set of species) in each recording

3. presence/absence of the target species (or set of species) in each time interval that the user specifies (by default 60 seconds)

In addition to the Excel file, for each sound file AviaNZ generates an annotation file of the automated detections that the user can open and review in either the main interface or using the 'Review Batch Results' option.

6 Training a species recogniser

6.1 Overview

One of the main features of AviaNZ is that you can train your own species recognisers for use in batch processing. You can also swap them between people.

There are three parts to training a filter:

1. Creating training labels

2. Running the training process

3. Testing the recogniser

Once these have been completed, the filter will be saved, and can then be used in the recognition process, for example in 'Batch Processing'.

Before we start considering the process of training, it is useful to understand how to interpret the outputs that AviaNZ gives you.

6.2 Some important concepts

The way that AviaNZ decides whether or not it has recognised a call correctly is by comparing it with human annotation. The training sound files that you provide, together with their annotations, are used to recognise the characteristics of the calls of a particular species. Comparing the human and machine outputs, there are four possible outcomes for each second of the recording:

		Human	
		Call	No Call
AviaNZ	Call	*True Positive (TP)*	*False Positive (FP)*
	No Call	*False Negative (FN)*	*True Negative (TN)*

True Positive (TP) AviaNZ and human agree that there was a call

True Negative (TN) AviaNZ and human agree that there was not a call

False Positive (FP) AviaNZ says that there was a call, but the human did not

False Negative (FN) AviaNZ did not detect a call that the human found

The counts of how many seconds of the recordings correspond to each of these four quantities can be combined to produce a variety of measures of accuracy. Two of them are particularly important for us:

Specificity (True Negative Rate) $= \frac{TN}{(TN+FP)}$ (number correctly labelled as negative / actual number of negative examples)

Recall (Sensitivity or True Positive Rate) $= \frac{TP}{(TP+FN)}$ (number correctly labelled as positive / actual number of positive examples)

False Positive Rate $= 1 - \text{Specificity} = \frac{FP}{(TN+FP)}$ (number incorrectly labelled as positive / actual number of negative examples)

Precision $= \frac{TP}{(TP+FP)}$ (number correctly labelled as positive / number labelled as positive)

Ideally both recall and precision would be close to 100%, but it is hard to do both at once. For this first part of the recognition process we aim to detect as many of the calls as possible (high recall), which can lead to lower precision.

We also compare the True Positive Rate and False Positive Rates to assist in parameter setting, as we shall see shortly.

6.3 Labelling

In order to start the training process, you need to start by performing some manual labelling of calls of your target species in the Manual Interface.

Preparation Select some sound files for training and testing.

- The selection of sound files for training, and the careful labelling of the calls within those sound files, has the most potential for making a good recogniser.
- Make new folders on your computer, for training and testing files, and copy the sound files into them.
- You should try to pick a set of sound files that display all of the call variations of that particular species.
- If the species shows geographical variation you should also pick them from across that range (or name the recogniser so that the geographical specialisation is clear).
- Ideally you will need a few (5–30) examples of each type of call that the species makes.
- The sound files can have noise in them, although preferably not too much.
- It is helpful if you have both loud and quiet calls.
- Ideally you should avoid recordings where there are simultaneous calls from other species in the same frequency range.
- It is also a very good idea to have a set of different files for testing.
- Just like the training data, testing data should represent the real nature of field recordings that you are going to process with this filter, and also include all the call types.

Labelling Once you have a few sound files, open each file in the 'Manual Processing' interface of AviaNZ and manually label every call from that species with the species name

- For each call that you want the recogniser to identify, drag a box around the call reasonably accurately, and label it with the name of the species.
- Label every call of that species, loud and quiet, except, for the training set, those that are nearly impossible to see.
- For birds with calls (i.e., sets of syllables in a sequence), mark the whole call of one bird with a single box.
- Include any harmonics that are visible.
- If you have more than one bird of the species calling, mark them both, using separate boxes.
- If the calls overlap, the boxes can too.
- You do not need to label calls of any other species of bird, unless you also want to train filters for them.
- Do the same process for both the training and testing folders.

6.4 Training a wavelet recogniser

Wavelet recogniser is the initial form of a recogniser in AviaNZ. All automated recognisers should start with training a wavelet recogniser and then later on it is extendable to have a CNN post processor, particularly to improve the precision of the automated call detection (Section 6.5).

Follow the steps below to train a wavelet recogniser:

- To begin training a recogniser, select 'Train an automated recogniser' → 'Train a wavelet recogniser' from the Recognisers menu.

- Then follow the instructions to select the folder, and the name of the species:
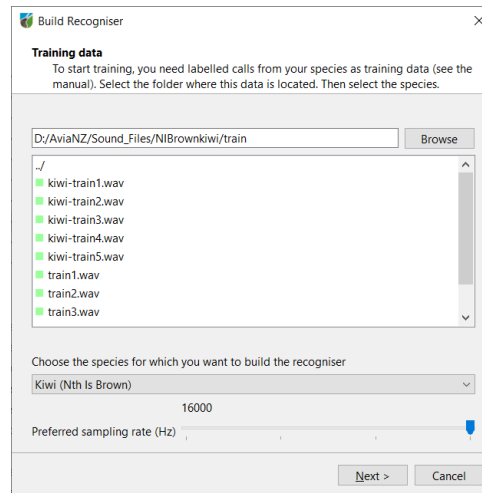
- AviaNZ will confirm these choices before starting the training process.

- There are a few places during the process where you can modify choices that AviaNZ has made.

- One is on the first page, where it says 'Preferred sampling rate (Hz)'. If you don't know what these things mean, you can ignore them, but for experienced users, they are places where choices that AviaNZ has made might be improved upon.

- Once the initial choices have been made, press the 'Cluster' button, and AviaNZ will try to cluster your labelled training segments into groups of similar sounds. It will show the outputs of this in this interface:
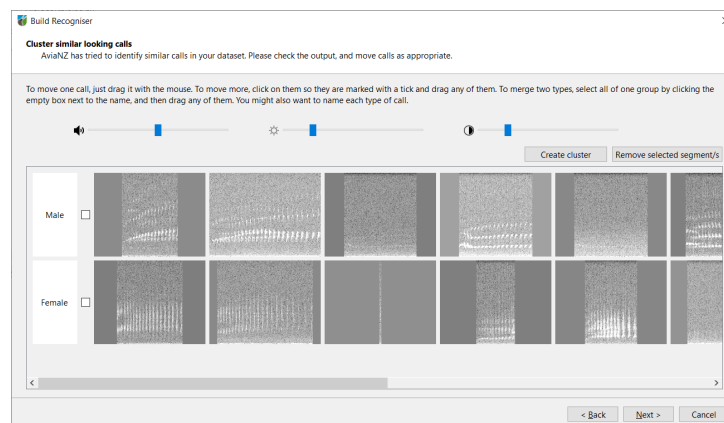
- We will improve this clustering over time, but at the moment it does make errors.

- There are two ways that you can improve the recogniser that is made:
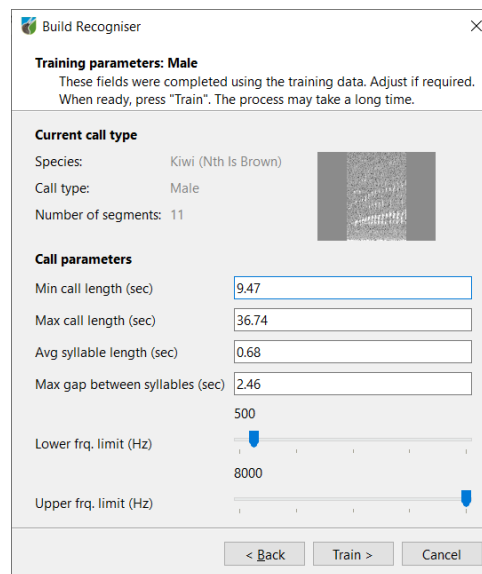
  Give names to the clusters It is intended that the clusters each represent different call types or sex of caller for some species. The default names are meaningless, since AviaNZ doesn't know about the species in advance, so it is normally useful to give a meaningful name by clicking on the name and typing a new one.

  Correct any errors by moving the spectrograms between different clusters as appropriate.
    - To move a single spectrogram, just drag it to the correct cluster
    - To move a whole group, click on each one (so it is marked with a tick) and then drag any of them to the correct cluster
    - To move some to a new cluster, click on them, and then click on the 'Create cluster' button.
    - To select all of the calls in a cluster, use the small tick box next to the name of the cluster
    - You can play the calls by clicking on the top-left corner of them.

- AviaNZ will now work through each cluster, and train an individual recogniser for that kind of call.

18

- It will show you a variety of parameter settings:

- If you don't know what they are, just ignore them.

- The program will then search for good ways to represent those calls, which can take a while, and then show a plot of error rates for different parameter settings.

- This is known as an ROC curve, and it is a plot of the recall (True Positive Rate) against the False Positive Rate for different parameter settings. It looks like these examples:
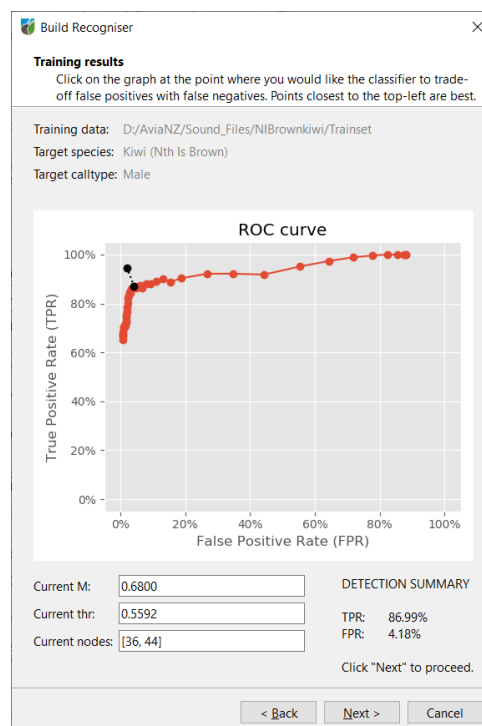
  - The perfect recogniser would be in the top-left corner of the graph.
  - Points lower down miss examples of the bird calling, while calls further to the right provide more false detections, i.e.., think the bird is calling when it is not.
  - You need to choose the compromise between these two that you are prepared to accept: more work in reviewing to get rid of false positives, or accepting that the software has not detected every call.

- To make the choice, click near the point that you think is the best compromise on the curve, which AviaNZ will show as a black dot (you can see it in the picture above).
- You don't have to click on it exactly, AviaNZ will show you which point is closest to where you clicked.
- You need to do this for each call type.

- In the Post-processing window, there is the potential to make another practical choice, which is whether or not to use the Fundamental Frequency calculation as part of the filter:
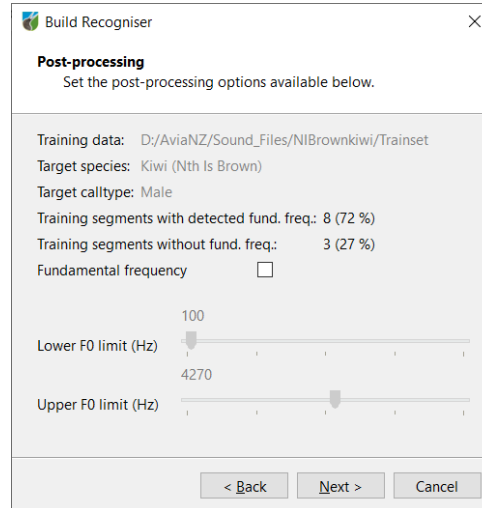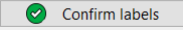

- For some calls, birds can produce a base note and then harmonics above that, and the fundamental frequency calculation tries to find the base note.
- Where this works it can be helpful, so if you compare the number of training segments where the software found a fundamental frequency with those where it didn't, you can see whether or not it makes sense to use it.
- You can also see the range of frequencies that were identified, so if they look wrong, don't use it.
- If you don't know, AviaNZ will try to work it out for itself.

- Following this, you should save the recogniser you have trained. It is then a very good idea to test it.

6.5 Extending a recogniser with CNN

A wavelet recogniser usually has good specificity which means that it can narrow down the search space to a greater extent. However, the common problem is still having to review lots of segments while the maority being represented by false positives. Therefore, we treat the wavelet recogniser as pre-processor and allow you to train an AI model – a Convolutional Neural Network (CNN) – on top of the wavelet recogniser. This combination allows you to improve the precision of the automated call detection.

Preparation Before proceeding to the CNN training, we recomend you to compile a resonable dataset for the CNN training following the steps below:

- Make a new folder, select and copy some field recordings into this folder. The selected recordings should represent the real nature of field recordings that you are going to process with this recogniser, e.g. consider various background noise, call variations etc. If you have lots of recordings it would be a good idea to select a random sample of managable size, instead of picking one recorder/location.
- Then Batch Process the selected recordings with the target wavelet recogniser.
- Following that you should Batch Review the results in the All Species mode.

- In addition to deleting any false positives as you would do in the normal form of review, here we want you to correct call types where necessary. This is bacause it's important that you show a correct set of examples from each call type of the target species. Use the little $\mathrm{SP_{CT}}$ button on the right-bottom corner of review page to see the wavelet suggested call type for each segment. If the recogniser suggests a wrong call type make sure you correct it.

- It is also important that you don't show unnecesarily long segments to the CNN: when a call only occupys a part of a segment. Instead of confirming such segments you have to mark them with ⑦ and proceed. At the end of the review open the folder in Manual Processing mode and navigate to the files you marked with the question mark. You can easily do that with the help of coloured file list and then use the overview to locate the target segment. Manually adjust the start/end points of the segment and press ✅ Confirm labels to confirm your choise.

- In addition to the batch reviewed recordings, you can also input manually annotated recordings to the CNN training, e.g. the training data set that you used for the wavelet recogniser training. If you plan to use some other manually annotated recordings (not the train data in wavelet training), make sure to annotate both species and call type levels. You can use $\mathrm{SP_{CT}}$ button from the Controls area in the Manual Processing mode to assign call types.

Training Follow the steps below to train your CNN:

- Select 'Train an automated recogniser' → 'Extend a wavelet recogniser with CNN' from the Recognisers menu.

- Then follow the instructions to select the folder/s with training data: Use appropriate 'Browse' buttons to load your manually annotated recordings and Batch Processed and Batch Reviewed recordings. Select the the name of the recogniser to extend with CNN from the list.
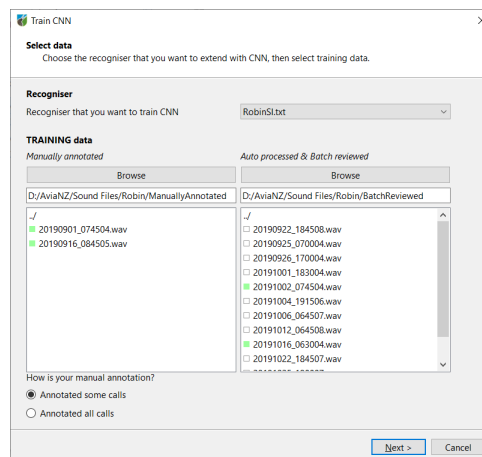
- AviaNZ will confirm these choices before starting the training process. And also will inform you when it requires more training data.

- When you proceed, AviaNZ will suggest a fixed length of a call/segment to feed to the CNN, and you can choose a sensible length for the species. The idea here is that the length should be long enough to see the syllable patterns on the spectrogram to differentiate target calls from noise, but not too long.

- At the end of the training, AviaNZ will show you ROC curves for each call type and ask you to hand pick a point that represents your compromise between true positive rate and false positive rate. However, if you want AviaNZ to look after those, remember to tick the box. Then AviaNZ will choose the best point in each ROC curve and show you a summary.
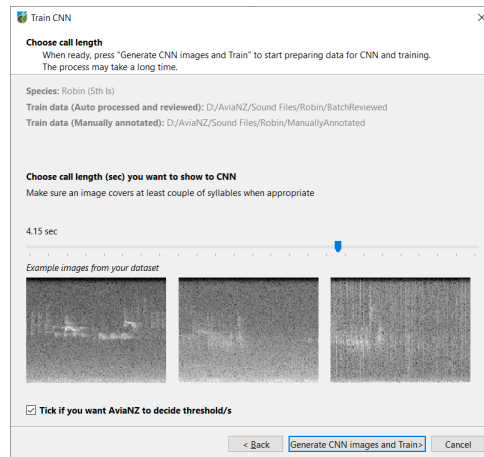
- It normally takes a long time to finish the training. Therefore please be aware of this before you hit the button 'Generate CNN images and Train'.

- Then if you choose to show ROC curves, ROC curves for call types will be displayed and you will be asked to choose a point. Otherwise, AviaNZ will automatically pick the best point on the curve.

- Following the summary of the CNN recogniser, save the recogniser you have trained. There are two options: you can choose to update the current recogniser or save it as a new recogniser. E.g. if you already got a CNN module in the current recogniser you may want to keep it and save the new recogniser to a new file.

- It is then a very good idea to test it on a seperate set of recordings.

6.6   Testing a recogniser

- You should always test a recogniser, and use different files than the ones you trained it on.

- If you have prepared testing data already, you can do this straight away.

- Otherwise save the recogniser and then test at a later date by using 'Test a recogniser' in the Recognisers menu.

- You should also test a recogniser that you receive from somebody else (for example, from our webpage) before relying on it.

- Testing should use different sound files to the ones used for training.

- It runs the recogniser over the files, and compares the results to the human annotations using the same error metrics that were defined in Section 6.2. It gives a better indication of how well the recogniser will work in practice.

6.7   Recogniser management

You can rename, import, and export recognisers by using 'Manage recognisers' in the Recognisers menu. If you have made recognisers for sounds that you think will be useful for other people, please upload them to our AviaNZ webpage.