

# The AviaNZ Bioacoustic Analysis Software (v 3.2-BirdNET)

The AviaNZ Team + Florian Meerheim

March 2024

This is the user manual for ▲ version 3.2 of the AviaNZ program with the BirdNET feature. There are several changes to this version of the program, of which the most obvious are (i) the ability to classify your recordings with the BirdNET classifiers, and (ii) several features to ease the analysis of the BirdNET results within the manual processing mode. These changes were introduced by Florian Meerheim. The purpose of this manual is to help you get started with using AviaNZ. There are also some videos for version 3.0 and an FAQ available at <http://www.avianz.net>. Furthermore a 'Cheat Sheet' is provided which you can find in the same directory as this manual or within the program under Help.

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.

## Contents

<b>1</b>	<b>Installation</b>	<b>3</b>
<b>2</b>	<b>Getting Started</b>	<b>3</b>
<b>3</b>	<b>The <i>Manual Processing</i> mode</b>	<b>4</b>
3.1	First Steps . . . . .	4
3.1.1	Cheat Sheet . . . . .	6
3.1.2	Navigate through files . . . . .	6
3.2	Spectrogram . . . . .	8
3.3	Manual Labelling (Audio Files) . . . . .	8
3.4	Manual Labelling (Bat Files) . . . . .	12
3.5	Controls . . . . .	13
3.6	Menu Options . . . . .	14
3.6.1	<i>File Menu</i> . . . . .	14
3.6.2	<i>Appearance Menu</i> . . . . .	15

3.6.3	<i>Actions Menu</i>	16
3.6.4	<i>Recognisers Menu</i>	17
3.6.5	<i>Utilities Menu</i>	17
3.6.6	<i>Help Menu</i>	18
3.7	Changing the Spectrogram Computation	18
3.8	Denoising	19
3.9	Interface settings	19
3.10	The BirdNET Workflow	22
3.10.1	The output of BirdNET	22
3.10.2	Classify recordings	23
3.10.3	Sorting the list of files	24
3.10.4	Filtering audio files	24
3.10.5	Manual review	24
3.10.6	Confirm/Change/Correct segments	25
3.10.7	Include previous/next recordings	25
3.10.8	Summarise results	26
<b>4</b>	<b>The <i>Batch Processing</i> mode</b>	<b>26</b>
<b>5</b>	<b>The <i>Review Batch Results</i> mode</b>	<b>28</b>
5.1	Quick Review	30
5.2	Review All Species	31
5.3	Outputs	32
<b>6</b>	<b>Training a species recogniser</b>	<b>33</b>
6.1	Overview	33
6.2	Some important concepts	33
6.3	Labelling	34
6.4	Training a wavelet recogniser	35
6.5	Extending a recogniser with CNN	40
6.6	Testing a recogniser	42

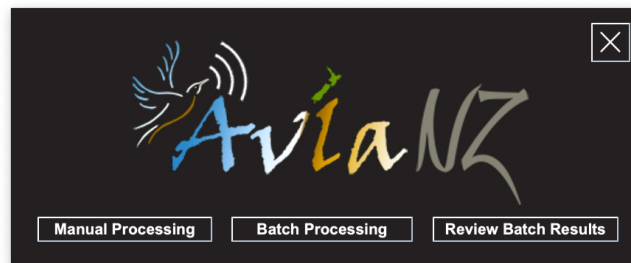
# 1 Installation

For Windows users there is a downloadable executable available on the github page <https://github.com/FloMee/AviaNZ/releases>. This should work for most users. If you plan on using the software under Linux or MacOS, follow the instructions in the README to install the software from source.

On Windows some files will be placed in your home directory, which may be somewhere like `C:/Users/username/AppData/Roaming/AviaNZ`. On Mac and Linux they will be in `~/ .avianz`.

# 2 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 3. It enables you to look at, listen to, and manually annotate individual audio files, as well as to train your own recognisers. Also the new features to classify recordings with BirdNET and other analysis options can be found here.
- The second option (*Batch Processing*) takes whole directories (and subdirectories) of audio files and automatically segments the calls of selected species you have recognisers for; see Section 4.
- You can view the output of the automatic segmentation using the third option (*Review Batch Results*, see Section 5), but if you want more context on them you could also use the first (*Manual Processing*) option.
- The software can also produce Excel files showing the results; these are described in more detail in Section 5.3.

### 3 The *Manual Processing* mode

Use this mode if you want to:

- Show and navigate through spectrograms of audio files
- Manually label bird calls in audio files
- Train your own recognisers
- Use the BirdNET classification models on your data
- Review the results of the BirdNET classification process

#### 3.1 First Steps

- 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 an audio file from a directory.

***Note:** AviaNZ can load audio files in wav format, or bat files in bmp format. Other audio files should be converted to wav format. AviaNZ assumes that audio files are recordings in mono (single microphone) sound. If there are other channels, only the first one is loaded. Very long recordings can be used, but it is generally easier if they are split into shorter files. We provide a splitter in the [Utilities Menu](#) but there are plenty of other options on the internet if you prefer.*

- The program will then ask you to give names for the operator and reviewer. These are useful to keep track of who has looked at different files.
- After selecting an audio file and setting the names the main interface (figure 1) shows up.

The name of the current file is shown in the title of the window (1). Below that there are four separate areas of the screen. Each has its name in a blue bar. They are:

**Overview (2–6)** This area shows a spectrogram of the first 5 minutes of the current audio file (labelled 2 in the figure). The blue section indicates the part you are looking at in the Spectrogram plot below. It can be moved by using **Left** and **Right** arrow keys, dragging the centre of the box, by clicking on the left and right arrow buttons (◀, ▶) next to the spectrogram (3) or by using the buttons in the bar below the spectrogram (4). It can be resized by dragging the ends of the box. Below the bar with buttons there are double arrow buttons (◀◀, ▶▶, labelled 5 and 6). With the buttons labelled 5 you can open the next 5 minutes page of the audio file, provided it exists. The other buttons (6) you can use to jump with the blue section to the next annotation segment.

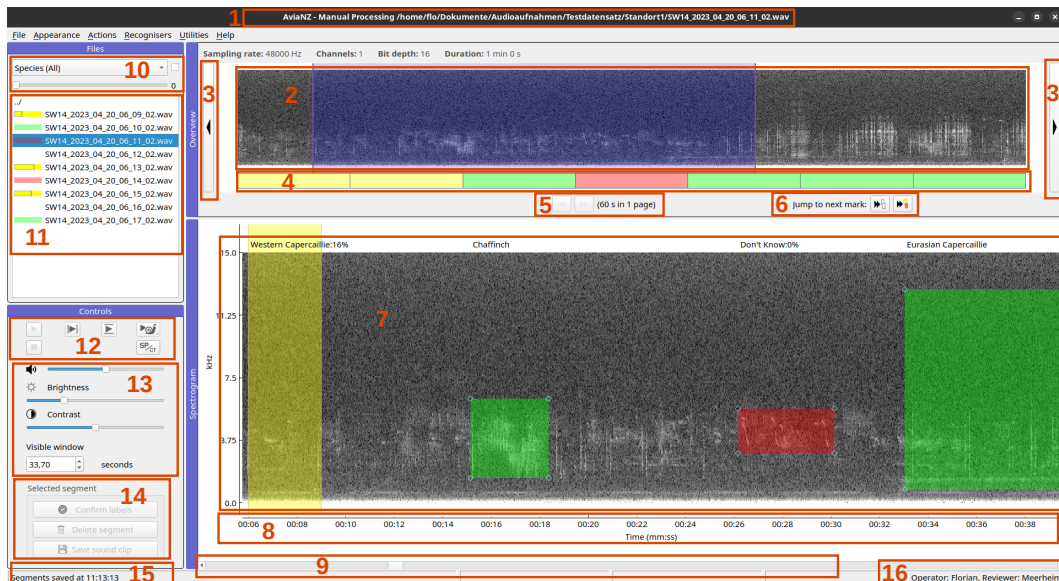


Figure 1: The interface of the *Manual Processing* mode with its different components.

**Spectrogram (7–9)** In this area the spectrogram of the blue section in the Overview area is shown. Here you can add and modify annotations. The axes for the image shows frequencies on the vertical axis and time on the horizontal axis (8). The time is true time if AviaNZ can read that from the file, and starts from time 0:00 otherwise. There is also a scroll bar to move through the current page (9). The way in which you annotate recordings is described in the section [Manual Labelling \(Audio Files\)](#).

**Files (10–11)** This is the list of audio files (11) in your current directory. Use the Up or Down arrow buttons on your keyboard to navigate through the file list. You can also double-click on a file to select and open it. The file icons next to the file names indicate the status of the audio file. This is described in detail in the section [Colour Codes](#). There are several filtering options (10) to reduce the file list. These are described in [Filtering audio files](#).

**Controls (12–14)** The buttons in this area are referring to the Spectrogram area. You can playback the audio file (12), modify the appearance of the plots (13) and process any segment that is currently selected (14). For more details, see [Controls](#).

There is an optional fifth area, which is the Amplitude plot. If you want to see it, select *Show amplitude plot* in the *Appearance* menu. Select the option again to hide it. You can also hide the *Files* and *Overview* area 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.

**Status Bar** At the very bottom of the screen there is an area that gives any status updates from the program (15), and on the right, information about the currently selected Reviewer and Operator (16).

- 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 with the name of the section in it. 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.
- If you want to move to a new directory, either use the ‘..’ option to navigate around your computer’s file system, or use the *Open audio 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*.

### 3.1.1 Cheat Sheet

You can find a Cheat Sheet with the most important navigation options, especially shortcuts in the *Help* menu under *Cheat Sheet*. Especially in the review process after the classification of your audio files with BirdNET it can help a lot.

### 3.1.2 Navigate through files

To quickly navigate through the list of files and the files themselves it is highly recommended to use the arrow keys on the keyboard: Use the **Up** and **Down** arrow keys to open the previous or next audio file in the file list. Use the **Left** and **Right** arrow keys to move backward and forward through the audio file.

But there are a lot of more options:

- To open an audio file:
  - Use the **Up** or **Down** arrow keys to open the next or the previous file in the file list.
  - Use *Open audio file* in the *File* menu if you want to open an audio file in another directory.
  - Double Click on a file in the list of files.
  - Double Click on a folder will open that folder; the ‘..’ option at the top of the list takes you up one directory.

**Note:** If some of the filter options in the *Files* area (10) are applied, some of the files might be hidden within the file list. In this case the **Up** and **Down** arrow keys apply only for the visible files. The shortcut **Alt + Up/Down** in contrast includes also the hidden files, so that you can still reach them.

- To move within a file:
  - Press the **Left** or **Right** arrow keys.
  - Drag the blue highlight in the overview spectrogram.
  - Click on any of the boxes in the bar below the overview spectrogram.
  - Click on the left or right arrows on the left and right of the overview spectrogram.
  - Drag the scroll bar below the spectrogram.

**Note:** According to the set page size (defaults to 5 minutes) in the [Interface settings](#) long files are split into several pages. To open the next or previous page of an audio file use the double arrow buttons (5) below the overview spectrogram or the shortcut **Shift + Left/Right**. More details in: [Moving Through Long Files](#).

- To change the amount of the file that you can see (the visible width):
  - Drag the ends of the blue highlighted section in the overview spectrogram.
  - Change the *Visible window width (seconds)* in the Controls dock 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 [Changing the Spectrogram Computation](#).

**Moving Through Long Files** According to the set page size (defaults to 5 minutes) in the [Interface settings](#) long files are split into several pages.

- To move to the next or previous page:
  - Use the double arrow buttons labelled (5) or
  - press **Shift + Left/Right** arrow keys.
- The program tells you which page you are currently on, and how many there are in total.
- You can change the page size and the amount of page overlap by choosing the relevant options in the [Interface settings](#) in the *Appearance* menu.
- The times on the axis below the spectrogram show locations in the full file.
- Note that operations like denoising and segmentation (described in the section [Actions Menu](#)) apply to the visible 5 minute portion of the file, not the whole file.

## 3.2 Spectrogram

- The main spectrogram plot (7) shows a section of the current audio file. The part you are looking at is highlighted in blue in the overview spectrogram (2).
- The axes of the spectrogram are time on the horizontal axis (8) 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 in the Controls section (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.
- To change the computation of the spectrogram you can use the *Change spectrogram parameters* from the *Appearance* menu. The options are described in detail in section 3.7.
- It is possible to display information about where the mouse is pointing on the spectrogram (time, frequency, energy value) below the axis by choosing *Show pointer details in spectrogram* from the *Appearance* menu, and to switch it off the same way.

## 3.3 Manual Labelling (Audio Files)

The aim of the labelling process is to add further information to the audio file. For this purpose AviaNZ uses labeled segments. A segment is a specific range in time and frequency within the spectrogram. Together with the information about time and frequency a species label (e.g. a bird species) and a call type label (for specific call types of the species) can be stored.

- To create segments, click and drag with the left mouse button on the Spectrogram. (The action of left and right mouse buttons can be swapped over in the *Interface settings*.)
- 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.
- 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 there if you have particular preferences.
- Segments are saved automatically, so that you can't lose your work.




## Creating and Labelling Segments

- There are three ways to create a segment. You change which of them to use 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.

**Note:** The original version of AviaNZ comes with a bird list for New Zealand. In the list the species names are given as 'Genus (Species)'; e.g. 'Kiwi (Little Spotted)'. Given it in that way a submenu for every genus is created with the species in it. Although this works great for common species names in English, it doesn't in other languages. It is possible to use other bird lists (that do not necessarily have to follow this formatting) using the *Interface settings*.

- 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'.
- The 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.

- By default the lists of bird names updates 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 create 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.
- You can choose whether or not to allow more than one species per segment. This can be useful for recordings of the dawn chorus and where you just want to label presence of a set of species. Use the *Default to multiple species* in the [Interface settings](#) to allow this. Then, 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.
- AviaNZ also allows you to specify a call type. This could be the difference between male and female vocalisations for e.g., kiwi, or different calls types such as 'more-pork' or 'cree' for ruru. You can either add specific call types manually or the choice of which call types are available is made in an automatic filter, which is covered in the section [Training a species recogniser](#). To swap between seeing the species and seeing the call type label use the SP/CT button () on the right of area 12 of the Controls panel or press the **Tab** key.

## 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. This will be either the species label or the call type label depending on the SP/CT button.
- Segments have blue diamonds at the corners, so that you can resize them, or you can move the whole segment by dragging it (this works 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 or pressing 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 or the shortcut **Ctrl + D**.

- 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:

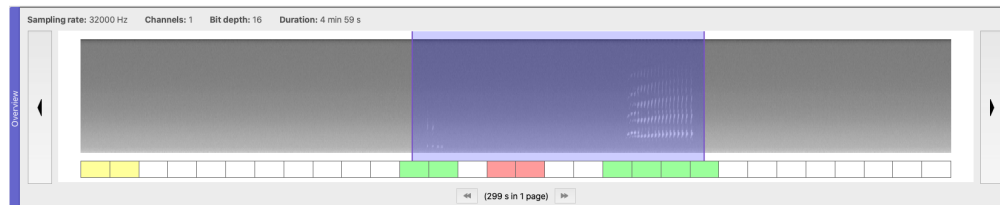
**Blue** This segment is currently selected. You can play it by pressing the buttons in the Controls area (see Section 3.5), or give it a new label by clicking on it again, or delete it.

**Green** This segment has been labelled with a bird name and a confidence value of 100. These are either manually created segments or the manually confirmed results of an automatic filter.

**Yellow** This segment has been labelled with a bird name and a confidence value between 0 and 100. These are either manually created segments with a question mark (by pressing **Ctrl** while dragging the box), or the row result of an automatic filter.

**Red** This segment has been labelled as ‘Don’t Know’. Its confidence value is 0.

- 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, or
  - Red** if there are ‘Don’t know’ segments, or
  - Yellow** if there are questionable segments, or
  - Green** if all the segments in that section are labelled with definite species.
- You can click on those boxes, and they will update the main spectrogram plot to show that section of the file. This is a useful way to move through the file quickly.
- The described color scheme also matches the colors of the file icon in the file list.
  - White** if there are no segments in the file or directory, or
  - Red** if there are ‘Don’t know’ segments, or
  - Yellow** if there are questionable segments, or

**Green** if all the segments in the file or directory are labelled with definite species.

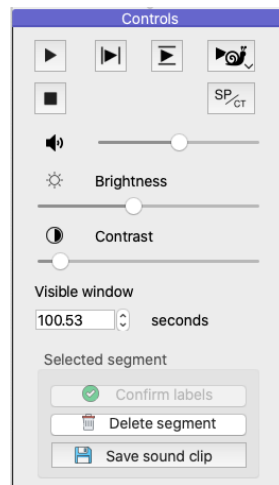
- Moreover, the grey box in the file icon indicates the maximum confidence value found in the segments of the file.

***Note:** The filtering options for the file list (10) also effect the file icon. If the segments in the file do not match the set conditions the file icons turn white. You can use this feature to quickly search for candidate recording after the classification process with BirdNET. The workflow with BirdNET is described in the section [The BirdNET Workflow](#)*

### 3.4 Manual Labelling (Bat Files)

- AviaNZ can also process bat files saved by DOC AR4 recorders (in .bmp format).
- When a bat file is loaded, the full length of the file will be shown, and the frequency range expanded up to 88 kHz. Guidelines will also be drawn on the screen to match the ones shown in DOC's BatSearch program.
- AviaNZ automatically makes a single segment that covers the detection for you. You can't see the segment, but you can see the 'Don't Know' label at the top of the spectrogram. Click anywhere on the spectrogram, and a menu will drop down so that you can label the species. If you view it as a 'possible' option then press the control key when you click to add a question mark to the label.
- You can delete the segment, but you can't add new ones or resize this one, the annotation is given to the whole recording.
- If you want to listen to a Bat call, there is an option to *Save audio file* in the *Actions* menu. This will create an audio file with the same name as the original file, but with a .wav extension. This can then be played as an audio file. We have slowed the recording down so that you can hear the bat sound in human audible range. Note that the quality of the audio file is not particularly high because of the way that the files are stored by the AR4.

### 3.5 Controls



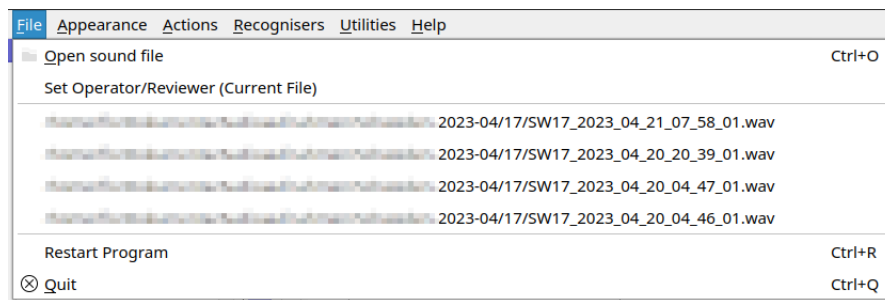
- The buttons at the top of the controls block allow you to play the sounds.
  - The top-left button is a normal *Play* button. The button turns into a *Pause* button while playing. Instead of the buttons you can start and pause the playback using the shortcut **Space**.
  - While the sound is playing, 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 three play buttons to the right of the stop button to play just that small segment.
  - The difference between these is:
    - \* the one on the left (▶) plays all the frequencies in the audio file (shortcut: **Ctrl + Space**)
    - \* the one in the middle (▢) plays 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 one on the right (with the snail: 🐌) lets you play the segment at different speeds. Hold the button down to change it to 2x (twice as fast) or 1/2 or 1/4 speed. Note that the pitch also changes, but it can be interesting or useful to hear all the components of a call.

- You can change the volume of playback using the slider below the buttons.
- The SP/CT button changes whether you control species labels or call type labels of the segments (shortcut: **Tab**).
- The *Brightness* and *Contrast* sliders change the appearance of the spectrograms (2, 7), helping to see more sounds.
- The size of the *Visible window* controls how much of the overview spectrogram (2) appears in the main spectrogram (7).
- The *Confirm labels* button makes a segment label certain. This is useful to confirm labels that were made by an automatic recogniser, or that the user was previously unsure about.
- The *Delete segment* button removes any segment that is selected (blue colour).
- The *Save sound clip* button saves a selected segment as a short audio file.

## 3.6 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.

### 3.6.1 File Menu



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

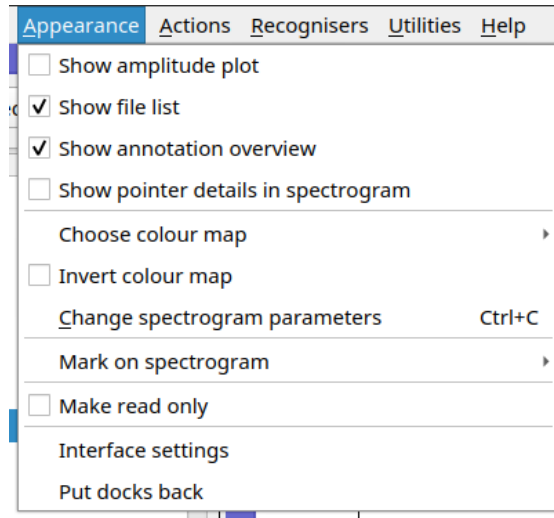
**Set operator/reviewer (Current File)** Enables you to change the operator or reviewer that were specified when you started AviaNZ. The change only applies to this file.

**List of recently used files** Makes it easy to revisit recent files.

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

**Quit** Does what it says.

### 3.6.2 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 for the spectrograms 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 the section [Changing the Spectrogram Computation](#) for more details.

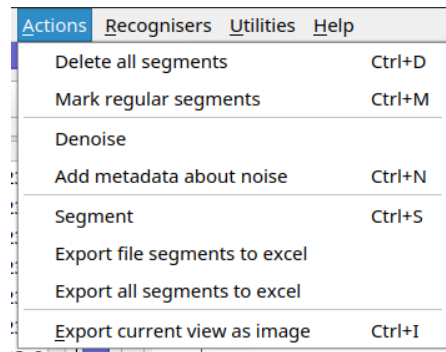
**Mark on spectrogram** There are some features that will help you spot bird calls or recognise them. We currently offer three options here: the fundamental frequency, spectral derivatives, and points of high energy. To select one, choose it from the menu; select it again to remove it. They can help 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 the section [Interface settings](#) for more details.

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

### 3.6.3 Actions Menu



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

**Delete all segments** Does what it says.

**Mark regular segments** Some users find it helpful to have an automatically generated segment box at regular intervals, for example to label all the birds in that segment under some statistical protocol. By default these boxes are made to be 15 seconds long, with one box every 300 seconds (5 minutes) through the file. These options can be changed using the *Check-Ignore protocol* area in the *Interface settings*. There is also an option there to simply mark those areas rather than putting a segment box in place.

**Denoise** Runs some programs that try to get rid of the noise in the audio file so that the birdcalls are easier to see. See Section 3.8 for more information.

**Add metadata about noise** Allows the user to specify if the audio 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 4) 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.

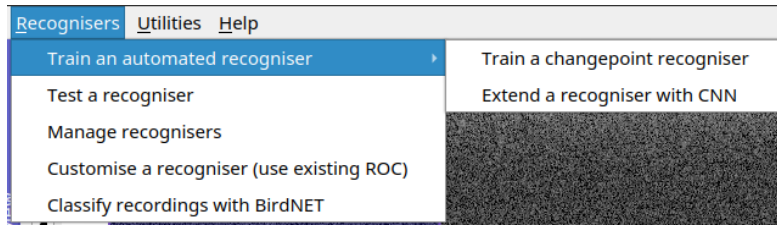
**Export file segments to Excel** This enables the user to save a summary of the annotation of the currently opened file into an Excel workbook (in the same location as the current audio file), as is described in section [Outputs](#).

**Export all segments to Excel** This option summarises all the annotations of the current directory (including subdirectories) into Excel workbooks as described in section [Outputs](#).

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



### 3.6.4 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 as simple as we can make it. It is described in section [Training a species recogniser](#). Please read that section before using this feature.

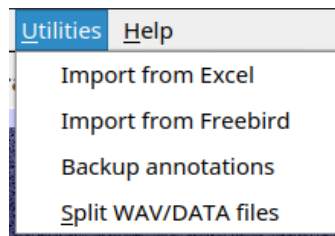
**Test a recogniser** Enables you to test a recogniser by choosing a folder to run it on. See section [Testing a recogniser](#).

**Manage recognisers** Lets you rename, export, or import new recognisers for different species. If you have made recognisers for sounds that you think will be useful for other people, please upload them to our AviaNZ webpage.

**Customize a recogniser (use existing ROC)**

**Classify recordings with BirdNET** Choosing this option a window will open where you can set the parameters for BirdNET to classify all your recordings in the current directory and in all subdirectories. It is described in more detail in the section [The BirdNET Workflow](#)

### 3.6.5 Utilities Menu



**Import from Excel** If you have made user annotations in other programs that you want to import, this option will help to do it. Normally you can export annotations in csv or excel format, and then load them into AviaNZ using this option.

**Import from Freebird** For the Freebird software, you can import user annotations directly.

**Backup annotations** If you want to copy a set of user annotations (for example to have them backed up, or to transfer them to another computer) without the wav files, but with the directory structure preserved, this option will do that.

**Split WAV/DATA files** DOC AR4 recorders produce 15 minute audio files, but many other recorders produce far longer recordings. This can mean that you spend a lot of time going forward and backward through spectrogram pages. This option lets you split the audio files into shorter pieces. If you have already processed them, so that they have segments included, there will also be transferred to the new audio files.

### 3.6.6 Help Menu

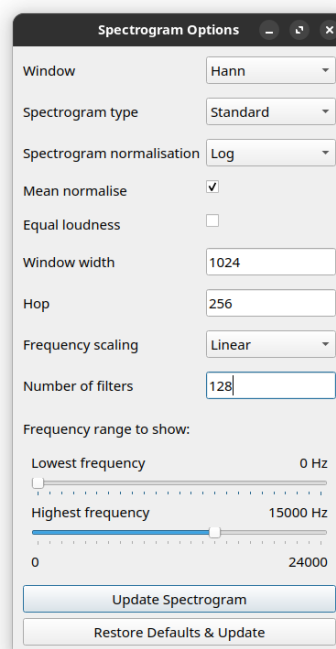
**Help** Gives access to this manual.

**Cheat sheet** Gives access to a cheat sheet with useful shortcuts and some information about the BirdNET classifier parameters.

**About** Shows you basic information about AviaNZ.

## 3.7 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 audio 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. It is computationally more efficient if these numbers are powers of 2 (such as 256, 512, 1024).
- 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.

### 3.8 Denoising

- When audio 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.

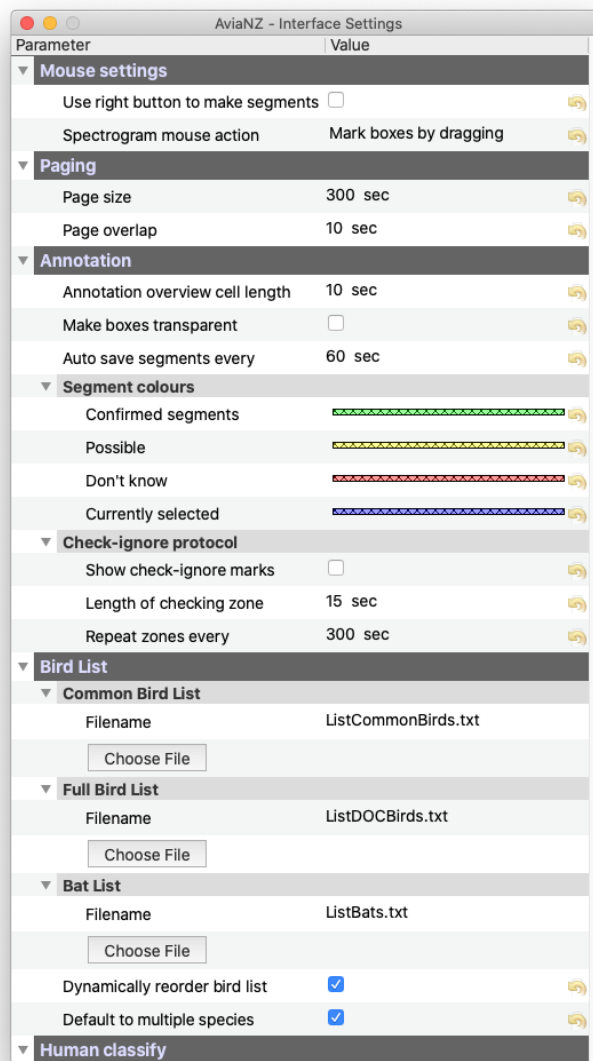
**Bandpass** Suppresses all frequencies outside a restricted range (which you specify) so that you can concentrate on the frequency range where calls that you are interested in can be seen.

**Butterworth bandpass** Another way to compute a restricted frequency range.

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

### 3.9 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 (described in detail in [Creating and Labelling Segments](#)).

**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, together with a bat list. 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.

**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.

### 3.10 The BirdNET Workflow

An overview of the workflow with BirdNET is shown on figure 2. It is divided into several steps, beginning with the classification process of the audio files, continuing with the filtering and reviewing of the detections coming to the summary of the results.

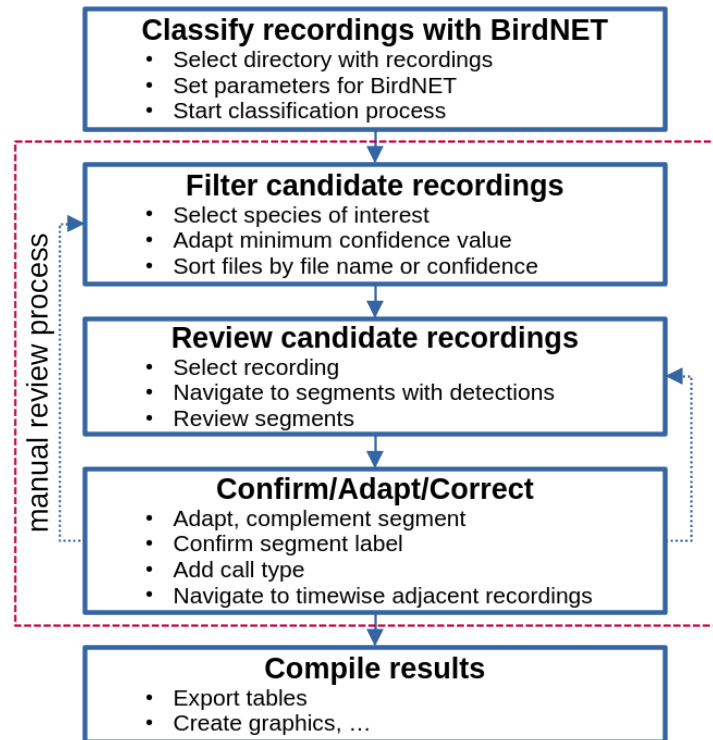


Figure 2: Overview on the classification and review process using the BirdNET classifiers.

#### 3.10.1 The output of BirdNET

Before the actual workflow is described the output of BirdNET will be introduced.

Internally BirdNET cuts your recordings into chunks of 3 seconds that overlap more or less depending on the set overlap parameter. For each of these chunks BirdNET produces a list of all classes BirdNET can distinguish (mainly bird species, hereafter species). To each of these species a confidence score ranging from 0 to 1 is attached. With setting the different parameters of BirdNET (for example: minimum confidence, latitude, longitude, week, custom species list) this list is reduced to a minimum of at most a few species.

For interpreting the results it is important to understand that the confidence scores are not calibrated. That leads to two important things:

- The confidence score is not a true probability or certainty value. A value of 0.8 does not mean, that on average the species can be found in 8 out of 10 segments labelled with this certainty value.

- The confidence scores can not be compared across different species. A value of 0.8 can have a completely different meaning for a pygmy owl than for a capercaillie.

### 3.10.2 Classify recordings

To start the analysis of your audio files with BirdNET you first have to specify different parameters. You can reach the input form by clicking on *Recognisers* → *Classify recordings with BirdNET* in the menu.

In this window the parameters for BirdNET-Lite or BirdNET-Analyzer can be set. The button *Show advanced settings* gives even more options. The meaning of the different parameters is explained in the cheat sheet that you can find over *Help* → *Cheat Sheet*.

If you are using BirdNET for the first time it is recommended to just set the geographical location with specifying *Latitude* and *Longitude* and provide the time of the year with the week parameter. The other parameters can stay with default values.

After the parameters are chosen, the analysis can be started by clicking on the *Analyze* button. That will classify all the recordings in the current directory including subdirectories. With that it is possible to, for example analyse all the recordings of different locations within one recording session. If you specifically want to use this possibility you can change to an upper directory via the `'../'` in the file list.

Depending on the number of recordings to classify, the CPU power and the number of threads the classification process might take minutes or several hours. Keep that in mind

while choosing your parameters. It might also be helpful to test your desired settings on a small test dataset to see if proper results are obtained.

### 3.10.3 Sorting the list of files

By default the list of files is sorted by their file names. Usually the recordings include a timestamp in the format Year-Month-Date-Hour-Minute-Second - or similar. Sorting the files by file name means sorting by time in these cases as well. In this mode moving to previous/following recordings in the file list also means moving to timewise adjacent recordings.

You can change the sorting behaviour of the list of files by using the option *Sort files by maximum confidence* in the menu under *Appearance*. This sorts the list of files descending by the highest confidence scores detected in the recording for the selected species. Moving to timewise adjacent recordings does not work in this mode anymore.

### 3.10.4 Filtering audio files

After the classification process finished you will see that the file icons changed. As described in the section [Colour Codes](#) the color indicates different states of the segments in the file. Moreover, the grey frame within the icon marks the maximum confidence score detected in the file.

The file icon always refers to the currently selected species. This can be selected after the classification process from the dropdown menu on top of the file list. The values after the species in this list indicate the maximum confidence score found for that species in the audio files of the current directory. If you also tick the box next to that dropdown menu, the file list will reduce to the files in which the selected species are found.

You can further reduce the list of files with changing the value of the slider below the dropdown menu. With that it is possible to exclude all the audio files with a maximum confidence value for the selected species below the selected value.

Use these options to systematically search for audio files with high confidence scores for the specified species (considered “candidate recordings”).

### 3.10.5 Manual review

If a first audio file from the file list is selected you can jump to the first segment labeled with the selected species either with pressing the button *Jump to next mark* below the overview spectrogram or by using the shortcut **Ctrl + Right**.

Use the play buttons to start the playback of the file. Alternatively you can use the shortcut **Space** or **Ctrl + Space** to playback the visible area in the spectrogram or just the selected segment respectively.

The goal of this step is to review all the segments within the file for the selected species.



### 3.10.6 Confirm/Change/Correct segments

If the segment is checked it should be confirmed, changed or corrected right after it. Only this way you can avoid to check the same segments and audio files again and again.

For that purpose different options exist:

- To confirm the label of a selected segment use the *Confirm labels* button in the *Controls* area or simply hit the **Enter** key.
- Use your mouse to drag single segments to change their position.
- Use your mouse and the diamonds in the corners of a selected segment to resize it.
- Use the *Delete segment* button or the **Delete** key to delete segments
- Use the *Delete all segment* option from the *Actions* menu or the shortcut **Ctrl + D** to delete every segment in the current file. This can be handy if you want to replace a lot of BirdNET detections in a file with just one segment.

***Note:** If there are several species attached to a single segment you can only confirm all the labels. In that case confirm all the labels first and deselect the “wrong” species by simply clicking with the right mouse button on the segment and deselecting the species from the opening menu.*

- If you can neither confirm nor refuse the label BirdNET has given to the segment it might be helpful to assign a custom ‘species’ label. Later, this label can be used to filter out recordings with these doubtful segments. The process how you can add custom labels is described in [Creating and Labelling Segments](#).
- Additionally to the species label AviaNZ provides the options to assign call type labels. You can change the type of the labels that are shown in the segments by pressing the *SP/CT* button in the controls area or with the **Tab** key. With right clicking on a selected segment a menu opens where you can select call types from a list or add a new call type for the specific species by choosing the *Add calltype* option.

After that process you can go on to the next file using the **Down** arrow key.

### 3.10.7 Include previous/next recordings

If a species was confirmed in an audio file it might be a good idea to look for further calls of the same species in the files that were recorded right before or after that. Because of the filtering process it might be that these recordings do not appear in the file list anymore. Nevertheless you can reach those recordings by pressing the **Alt** key together with the **Up** or **Down** keys when moving on to the previous or next audio file. For this option you have to sort the files by name, not by confidence score!

This process might be helpful to gain further ecologically interesting information about the species. Especially faint calls are still a challenge for recognizers like BirdNET.

### 3.10.8 Summarise results

If every candidate recording is reviewed the annotations can be exported to an Excel workbook. For that you can use the *Export all segments to excel* option from the *Actions* menu.

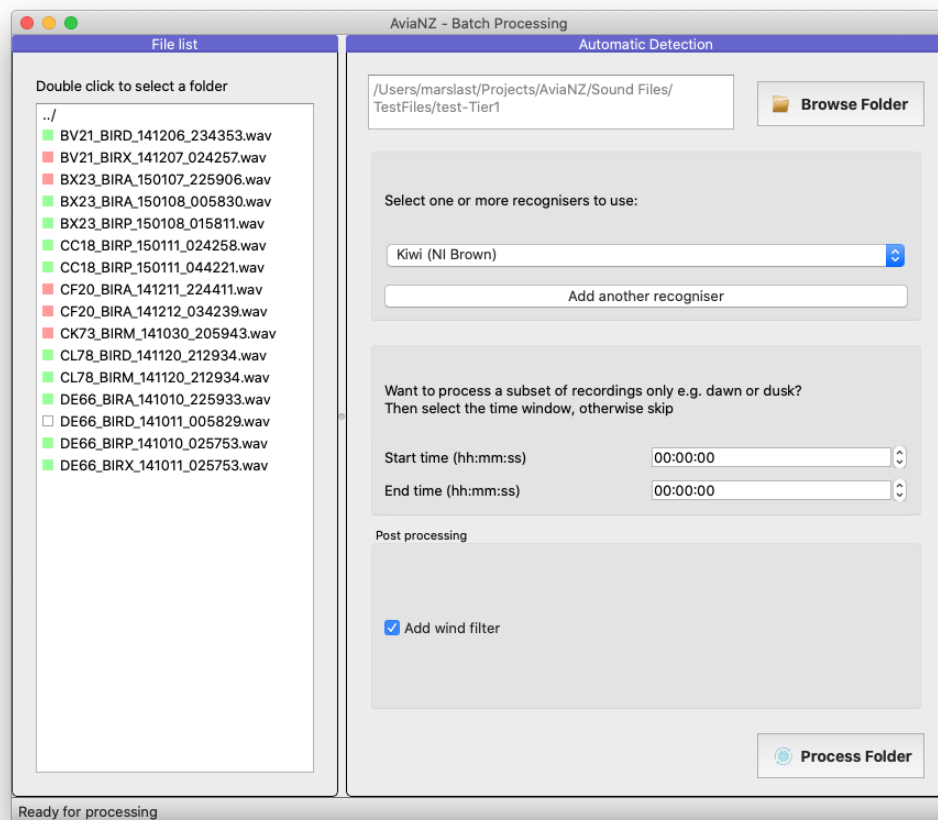
This option creates several Excel files in the current directory. The *DetectionSummary\_Any\_sound.xlsx* contains every segment of all the species within the current directory including subdirectories. Furthermore, individual files for every species are created with the pattern *DetectionSummary\_SpeciesName.xlsx*.

In these output files you can find a table in which for every segment the file, the start and end time of the segment as well as the minimum and maximum frequency and call type label are given. In the Excel file for 'any sound' a further column with the species label is attached.

You can use these exported data for example for further analysis or simply to create graphics.

## 4 The *Batch Processing* mode

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. After downloading all of the data from the SD cards into a folder on your computer, start AviaNZ.



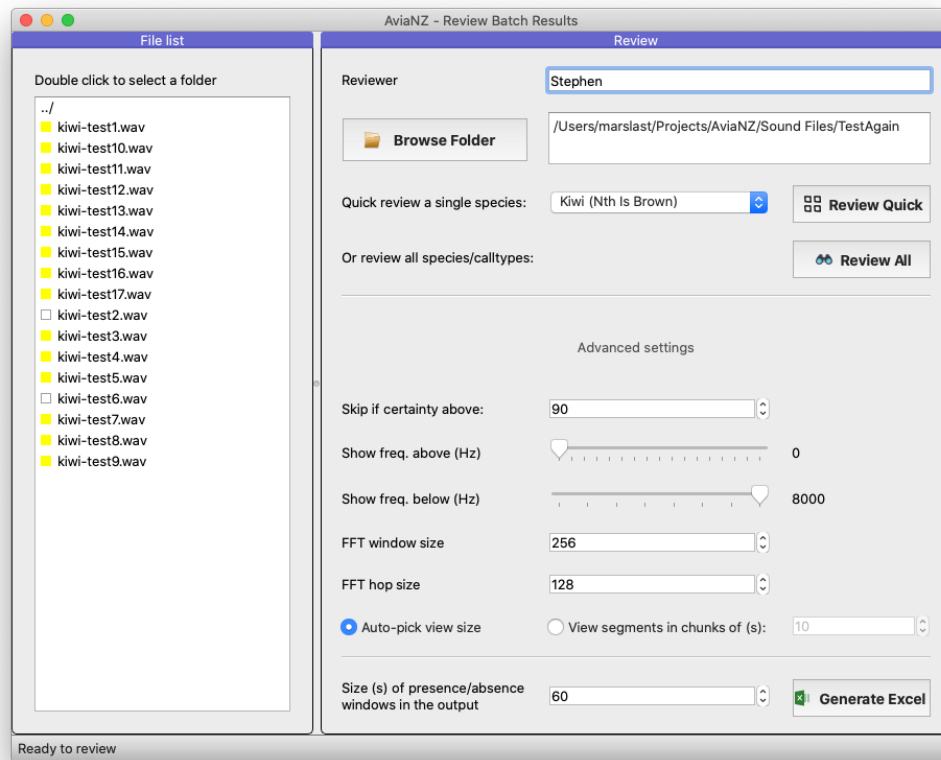
- Select *Batch Processing* from the starting window. You will see a screen like the one above.
- Navigate to a folder containing recordings to process.
- If this folder contains subfolders, the program will work through all of the folders inside the original one.
- Note that it deletes any previous annotations in your files. If you want to save them, use the manual mode and the *Utilities* → *Backup annotations* option.
- 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 5.3).
- You can choose one or more species-specific recognisers from the drop-down list to apply to the audio 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. This is useful if your data was collected in a windy place, but may miss some calls otherwise.
- 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.
- The *Any sounds (Intermittent sampling)* option shows regular segments based on the *Check-ignore protocol* settings in the *Interface setting* in Manual mode, which is 15 seconds every 5 minutes by default. This provides another way to perform some manual annotation quickly and easily.
- 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. If you have a lot of files it do this kind of processing 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.

## 5 The *Review Batch Results* mode

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 try not to have too many false negatives (missed calls) because they are harder to find later. However, this does mean that there are more false positives instead. We provide two interfaces for checking and correcting the results.

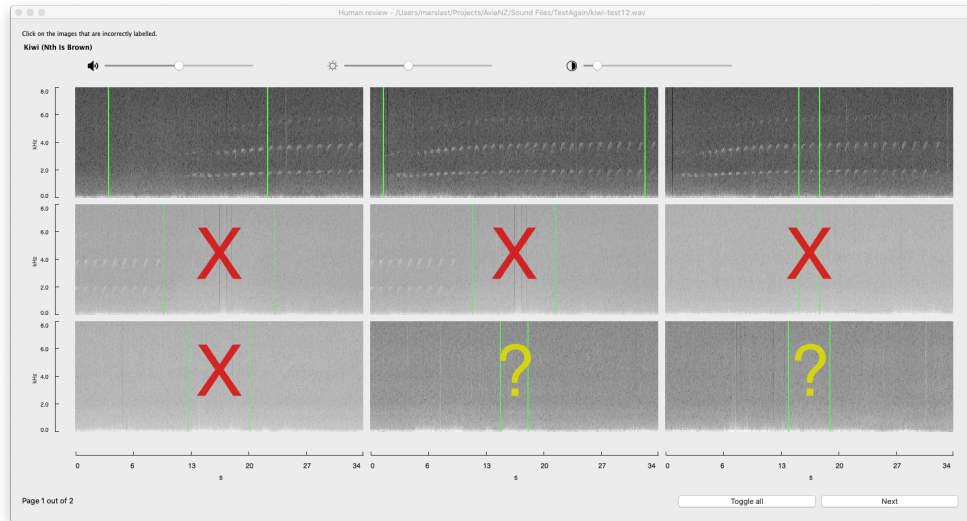
If you select *Review Batch Results* on the start screen then the following window will appear.



- Give a reviewer as in the Manual interface, then browse to a folder of previously processed files that you wish to review.
- *Quick review* is designed to let you process lots of files quickly to remove false positives. It shows only the spectrograms related to the species selected in the drop-down box, see Section 5.1.
- *Review all* is designed to let you see more detail about each segment, including potentially the call type as well as species. See Section 5.2.
- The option marked *Skip if certainty above* is important. AviaNZ assigns a confidence score to its classification of a bird call. User annotations, and automated annotations that the user has confirmed, are scored 100. Classifications by just a wavelet-based recogniser are scored 50, as are annotations with a ? on by the user. Neural-network based automated annotations have confidence assigned by the neural network. If you want to see all annotations you should set this parameter to 100 (the maximum value). If you wish to only see those that the system is not certain of, the default of 90 is suitable. As you set the value lower, more and more of the detections will be suppressed. Below 50 there will not be any segments to review.

- At the bottom of the window is the option to generate an Excel file of outputs; see Section 5.3.
- There are also a variety of options to change the appearance of the spectrogram in the Advanced Settings. These act in the same way as was described in the Manual interface section.

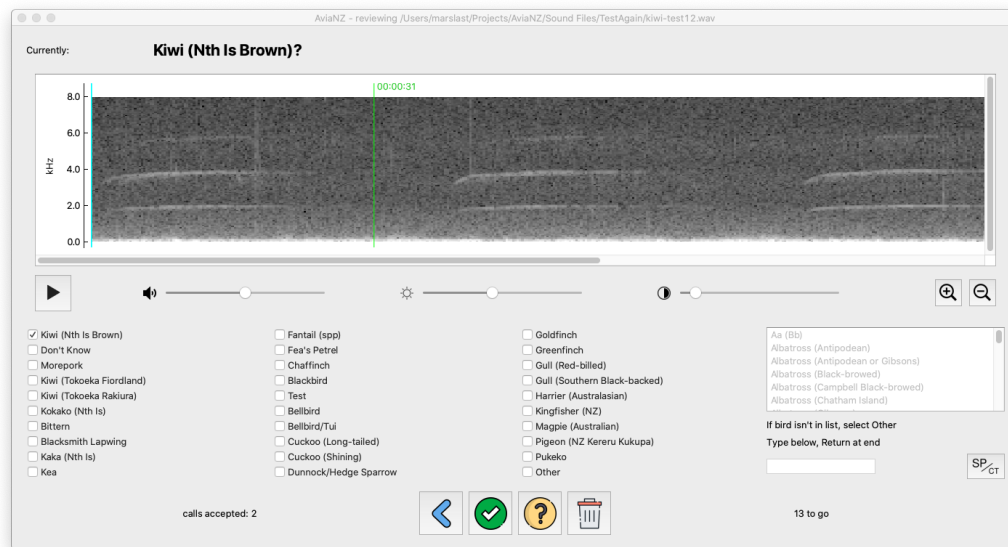
## 5.1 Quick Review




- The quick review mode is intended to let you delete false positives quickly.
- A set of spectrograms that have been labelled as your chosen species are shown. You can specify the size of these in the Batch Review interface above (by selecting *View segments in chunks of (s)*) or let AviaNZ choose it (*Auto-pick view size*).
- 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 (confidence 50), while those left plain will be confirmed as correct. Using the question mark option is useful to then review them further in either the All Species review (Section 5.2) or the Manual interface.

- You can change the brightness and contrast at the top of the screen.
- 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.

## 5.2 Review All Species



- The all species review is intended for a more in-depth review. You can also change the call type as well as the species using the  button.
- 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.
- If the spectrogram is too large then the window will scroll. You can use the  $+$  and  $-$  buttons (with a magnifying glass on) 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 are unsure, or want to mark the segment for further attention (for example, because it includes other sounds and you want to resize it) then press the question mark button (?). It is then easy to find in the Manual interface later.
- 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.

### 5.3 Outputs

The AviaNZ program aims to provide detailed and easy-to-review outputs. An Excel file (with the name ‘DetectionSummary’ followed by the 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 audio 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.

The process of training filters is fairly complex, and can be computationally expensive. The quality of the final filter depends on the training data used, the quality of the labelling that was performed, how noisy the files are, and how many other birds there are present.

AviaNZ has two levels of filters. The first is the wavelet filter. This does not need too much data to train, but will usually produce a lot of false positives. You can stop at this stage, but if you want to get higher accuracy, then you can train a second level filter, which refines the first by using a neural network. This requires a lot of data and training time. In both cases 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 using *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 audio 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	<i>True Positive (TP)</i>	<i>False Positive (FP)</i>
	No Call	<i>False Negative (FN)</i>	<i>True Negative (TN)</i>

**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, including:

**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)

Not all of these metrics are truly useful for birdsong recognition, mostly because there are usually far more seconds in a field recording without any bird calls in that there are with the calls, and so the true negatives swamp the calculations. The two of most interest are recall, which counts the percentage of the calls that were found by the program, and precision, which counts the percentage of the calls found that were correct. Ideally both recall and precision would be close to 100%, but there is a trade-off between them, so it is hard to do both at once. For the 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 audio files for training and testing.

- The selection of audio files for training, and the careful labelling of the calls within those audio files, has the most potential for making a good recogniser.
- Make new folders on your computer, for training and testing files, and copy the audio files into them.
- You should try to pick a set of audio 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 audio 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.
- 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 audio files, open each file in the *Manual Processing* interface of AviaNZ and manually label every call from that species with the species name.

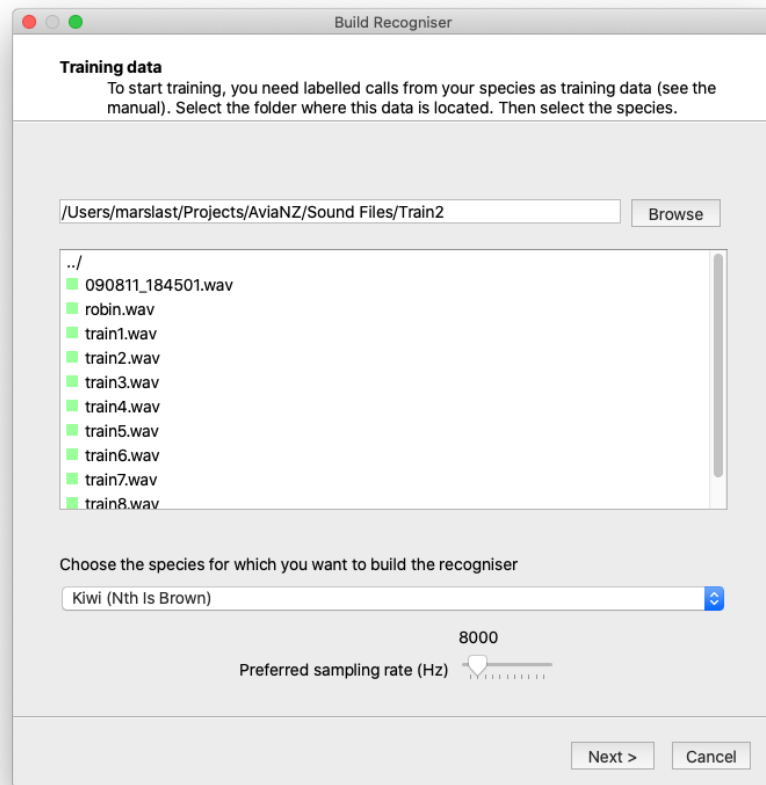
- For each call that you want the recogniser to identify, drag a box around the call reasonably accurately, and label it with the name of the species.
- Label every call of that species, loud and quiet, except, for the training set, those that are nearly impossible to see.
- For birds with songs or complex 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

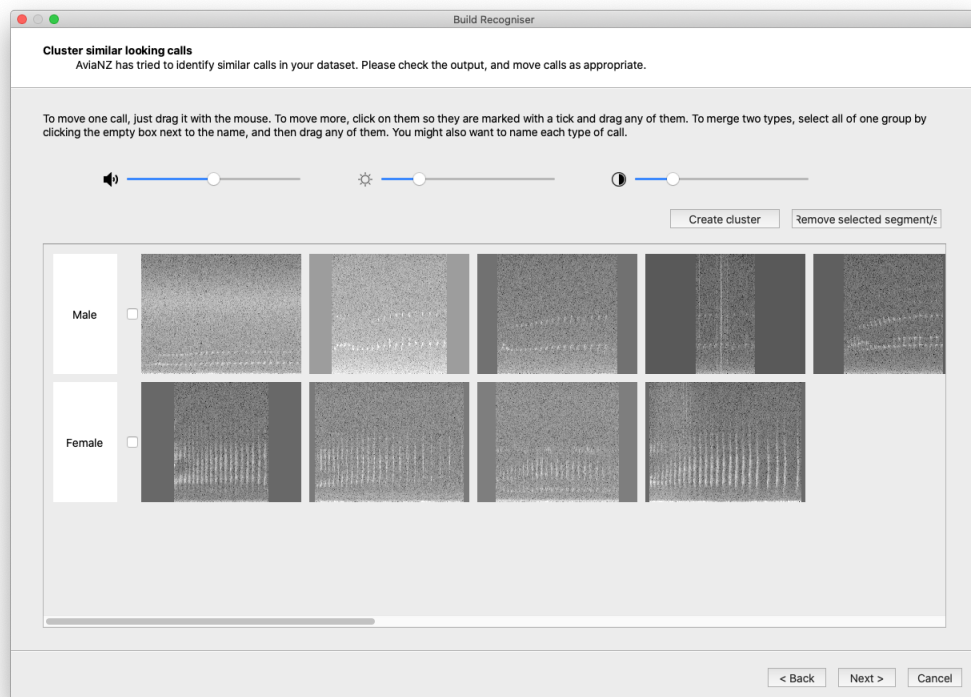
Wavelet recogniser is the initial form of a recogniser in AviaNZ. All automated recognisers should start with training a wavelet recogniser and then later on it can be extended with a neural network filter.

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 in the Manual interface.
- 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, there 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 (call types, such as male and female kiwi, or 'more-pork' and 'cree' ruru calls). It will show the outputs of this in this interface:



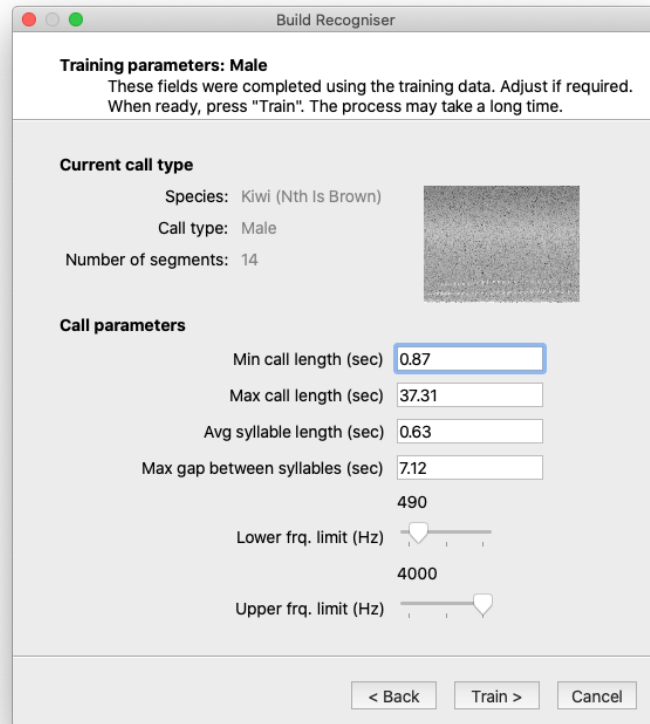
- We will improve this clustering over time, but at the moment it does make a lot of errors.
- There are two ways that you can improve the recogniser that is made:

**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.
- It will show you a variety of parameter settings:



**Build Recogniser**

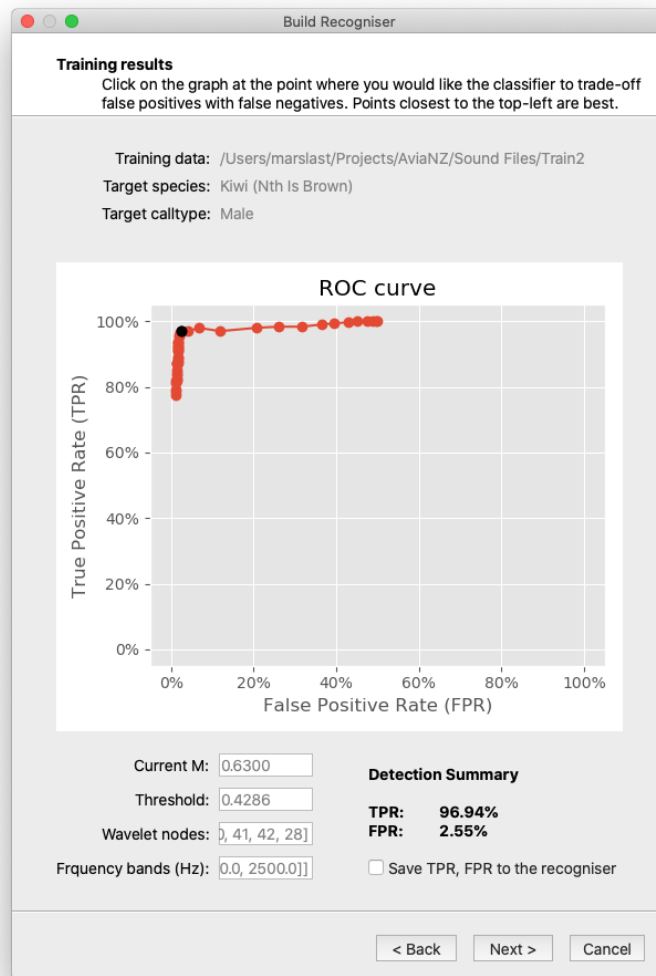
**Training parameters: Male**  
 These fields were completed using the training data. Adjust if required.  
 When ready, press "Train". The process may take a long time.

**Current call type**  
 Species: Kiwi (Nth Is Brown)  
 Call type: Male  
 Number of segments: 14

**Call parameters**

Min call length (sec)   
 Max call length (sec)   
 Avg syllable length (sec)   
 Max gap between syllables (sec)   
 490  
 Lower frq. limit (Hz)   
 4000  
 Upper frq. limit (Hz)

- 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:



- 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. If you are planning to train a neural network filter as well, choose to accept more false detections.
- 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.
- Following this, you choose a name to save the recogniser you have trained. It is then a very good idea to test it.


## 6.5 Extending a recogniser with CNN

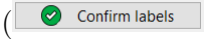
A wavelet recogniser will usually find most of your bird calls (it has good recall), but the price you pay for this is that it finds a lot of false positives (poor specificity). This leads to a lot of wrong segments that have to be reviewed. You can extend the filter in AviaNZ using a neural network known as a CNN (convolutional neural network). These are often used for image recognition. Training a neural network requires a lot of training data and computational power. The way that you perform the training is as follows:

**Preparation** You will need to compile a good collection of the calls of your species.

- 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 multiple recorders and sites, try and use examples from all of them. The more data you have here, the better. We have typically used enough data to have hundreds of examples of each call type.
- Once you have this data use the *Batch Process* option to apply the trained wavelet recogniser to the set of files. This will likely find the vast majority of the calls, but also a lot of false positives. It may also miss parts of calls, or make the calls too long.
- You now need to review these calls. We use a two-stage process for this.


**Remove false positives** Using either of the batch review options, delete any segments that do not contain your target bird call.

**Correct the segments** Using manual mode, correct the size and location of each segment box so that it covers the call, but only the call. Also check that the call type was labelled correctly and correct any errors. You can swap between seeing the species label and the call type label using the  button on the right of the Controls panel. You can also do this second part in the All Species Batch Review (Section 5.2).

Once you are sure that each label is correct, use the *Confirm labels* () button so that the segment turns green. Or at the end of manual correction of the box sizes, you can review the folder in Batch Review and confirm them to make it faster.

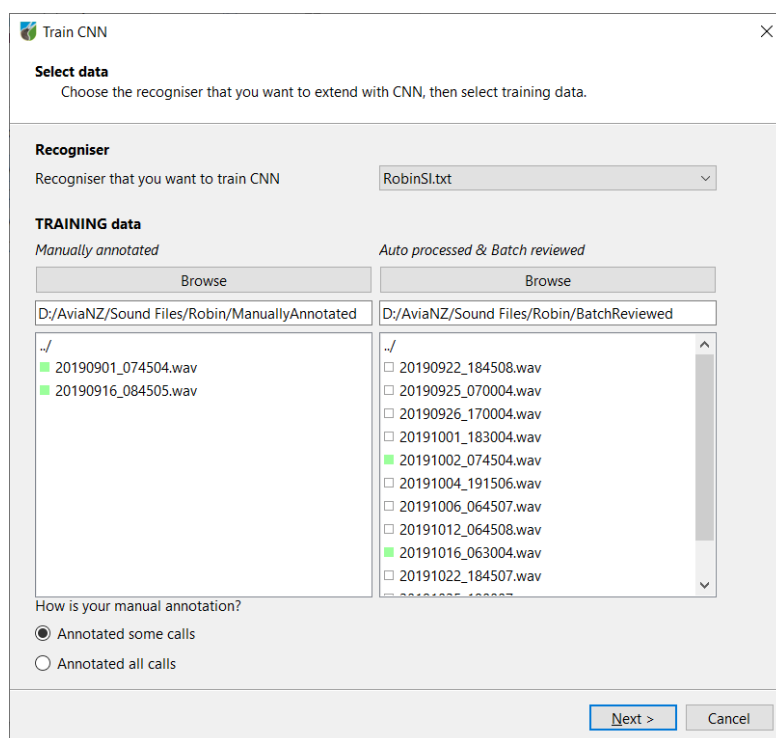
- In addition to the batch reviewed recordings, you can also input manually annotated recordings to the CNN training, such as the training data set that



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 it to include both species and call type. You can use  button from the Controls area in the Manual Interface to assign call types. Alternatively, you can annotate the calls to the species level and assign the call type during the review of the folder in the Batch Review All Species review (Section 5.2).

**Training** Follow the steps below to train your CNN:

- Select *Train an automated recogniser* → *Extend a wavelet recogniser with CNN* from the Recognisers menu.
- Follow the instructions to select the folder/s with training data. Use the appropriate *Browse* buttons to load your manually annotated recordings and the Batch Processed and Reviewed recordings. Select the the name of the wavelet recogniser that you will be extending from the list.



**Train CNN**

**Select data**  
Choose the recogniser that you want to extend with CNN, then select training data.

**Recogniser**  
Recogniser that you want to train CNN: RobinSI.txt

**TRAINING data**

Manually annotated	Auto processed & Batch reviewed
Browse	Browse
D:/AviaNZ/Sound Files/Robin/ManuallyAnnotated	D:/AviaNZ/Sound Files/Robin/BatchReviewed
<ul style="list-style-type: none"> <li>20190901_074504.wav</li> <li>20190916_084505.wav</li> </ul>	<ul style="list-style-type: none"> <li>20190922_184508.wav</li> <li>20190925_070004.wav</li> <li>20190926_170004.wav</li> <li>20191001_183004.wav</li> <li>20191002_074504.wav</li> <li>20191004_191506.wav</li> <li>20191006_064507.wav</li> <li>20191012_064508.wav</li> <li>20191016_063004.wav</li> <li>20191022_184507.wav</li> </ul>

How is your manual annotation?

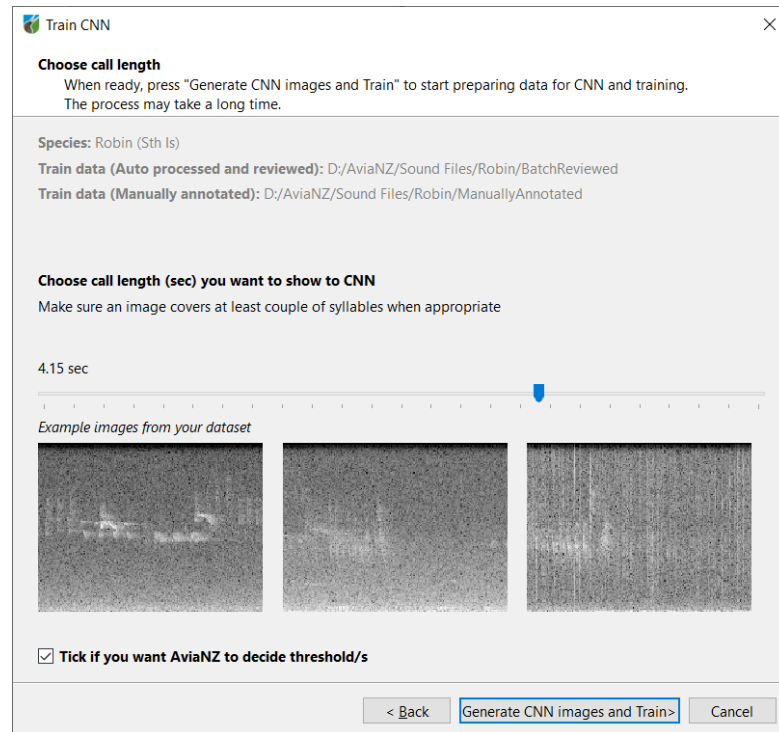
☒ Annotated some calls

☐ Annotated all calls

Next > Cancel

- AviaNZ will confirm these choices before starting the training process. It will also warn you if it cannot find enough data.
- AviaNZ will then suggest an appropriate length of call/segment to feed into the CNN. If the number looks wrong for your species, you can correct it. Ideally this length should be a bit longer than both a syllable of the call, and the gap between syllables.

- At the end of the training, AviaNZ will show you ROC curves for each call type and ask you to select a point that represents your compromise between True Positive Rate and False Positive Rate. If you select the relevant option then AviaNZ can do this automatically instead.



- It normally takes a long time to train the neural network.
- Following training, AviaNZ will show you a summary of the CNN recogniser, and you can choose to save it either as a new recogniser, or updating an old one. This means that you can train multiple recognisers for a species to see which works best.
- It is important to test your recogniser on some different data. If the results are not good enough, you will need to rerun the process with more training data and make sure that your labels are good.

## 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 of the Manual interface.

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