BASSA
BioAcoustics Superlet Scalogram Analyser

Table of Contents

Contact the Developer: b.jancovich@unsw.edu.au

Introduction

BASSA is a tool for time-frequency analysis of audio signals. It was designed to facilitate the visualisation of animal sounds with greater accuracy and detail than conventional short-time Fourier spectrogram methods (Jancovich, 2023). Fourier-based time-frequency visualisations are known as spectrograms, whereas BASSA produces “Scalograms”. Scalograms look like spectrograms but are computed using wavelet-based algorithms. The specific algorithm underlying the scalograms produced by BASSA is the “Superlet transform” also known as the SLT (Moca et al., 2021).

The SLT is appropriate for time-frequency analysis of any animal sound, though it excels most when applied to complex, low frequency sounds. It is particularly adept at visualising the fine temporal details in glottal pulses, amplitude modulations, noisy, and non-linear sounds, while also retaining high resolution frequency details. This makes BASSA especially useful in the study of vocal production methods, vocal anatomy and vocal biomechanics. While the suggested applications lie in bioacoustics, BASSA is useful for time-frequency analysis of any time-series data, from human speech, to brainwaves, to seismographic data.

One of the SLT’s greatest strengths is that it requires minimal tuning to produce optimal visualisations for any arbitrary signal. Fourier-based spectrograms require careful tuning of their window size, type, and overlap to produce useful visualisations. Those parameters must be optimised for each signal of interest, and depend on the goals of the analysis, which might be, for example, to prioritise temporal details of a sound. This is not the case for the SLT, which can produce accurate visualisations for a wide range of signals, with no parametric tuning. The user interface of BASSA has been designed with this in mind, and while tuning of the SLT is possible, dynamic optimisation of default algorithm values means manual tuning is not strictly necessary. For deeper explanation of SLT parameters, see 5.4. Superlet Parameters: and the original “Superlets” publication: Time-frequency super-resolution with superlets.

BASSA was designed as part of Cetacean research at the MammalLab, in the Evolution and Ecology Research Center, University of New South Wales, Sydney, Australia.

Contact the Developer: b.jancovich@unsw.edu.au

1. System Requirements & Installation

- Operating System: Windows 11 (64-bit)
- Processor: Intel or AMD x64 processor with four logical cores
- Memory (RAM): 16 GB or more
- Storage: At least 2 GB of free disk space
- Display: 1366 x 768 or higher resolution monitor

Earlier versions of Windows may be compatible but are currently untested. Some functionality may be inoperable or unstable on earlier operating systems. Virtual machines also have not been tested.

Other requirements:

- MATLAB Compiler Runtime (MCR) version 9.9. The BASSA installer will download and install this automatically if not already installed.
- Microsoft .NET Framework 4.8 or later.

It is recommended to have a stable internet connection to download and install the application, as well as any necessary updates. Please note that some anti-virus or firewall software may interfere with the installation or functioning of the MCR.

2. Supported Audio Import File Formats

Note that BASSA currently supports only single channel audio. If a stereo or multi-channel audio file is loaded, a dialog box gives the user the option to select which channel to work on.

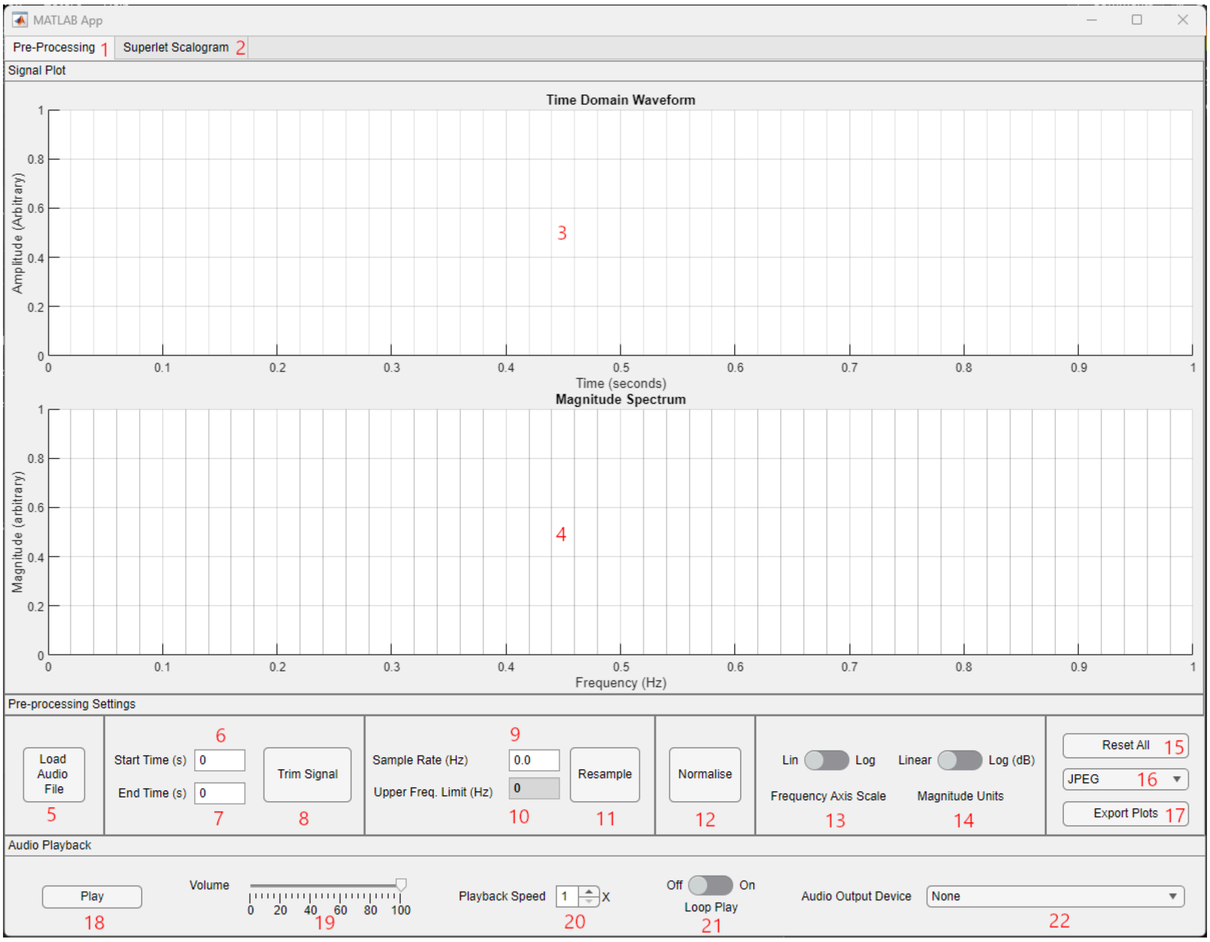
- AIFC (.aifc)
- AIFF (.aiff, .aif)
- AU (.au)
- FLAC (.flac)
- OGG (.ogg)
- OPUS (.opus)
- WAVE (.wav)
- MP3 (.mp3)
- MPEG-4 AAC (.m4a, .mp4)

3. Supported Image Export File Formats

- JPEG (.jpg)
- PNG (.png)
- TIFF (.tif)
- PDF (.pdf)
- EPS (.eps)

Contact the Developer: b.jancovich@unsw.edu.au

4. The User Interface



4.1. The Tab Bar

1. Pre-processing tab select button.
2. Superlet scalogram tab select button.

4.2. The Pre-Processing Tab

3. Waveform plot area. This shows the signal that will be analysed by the SLT in the time domain.
4. Magnitude spectrum plot area. This shows the signal that will be analysed by the SLT in the frequency domain.
5. Load Audio File Button. Opens a file load dialog box.
6. Edit audio start time (in seconds). Disabled after trimming, re-enabled by “Reset All”.
7. Edit audio end time (in seconds). Disabled after trimming, re-enabled by “Reset All”.
8. Trim audio signal to be analysed by the SLT, according to the new start and end times. Changes made to the signal start and end times are reflected in the signal sent to the audio playback panel. Disabled after trimming, re-enabled by “Reset All”.
9. Edit the audio signal’s sample rate. The lower the sample rate of the audio, the faster the SLT analysis will be. Sample rate determines the highest frequency that can be analysed. Disabled after resampling, re-enabled by “Reset All”.
10. Read-only: The new highest frequency that will be analysed as per the new sampling rate.
11. Resample the audio signal to be analysed by the SLT, according to the new sample rate. Changes to the sample rate do <u>not</u> affect the audio signal sent to the playback panel. Disabled after resampling, re-enabled by “Reset All”.
12. Normalise the audio signal to be analysed by the SLT, and the

Contact the Developer: b.jancovich@unsw.edu.au

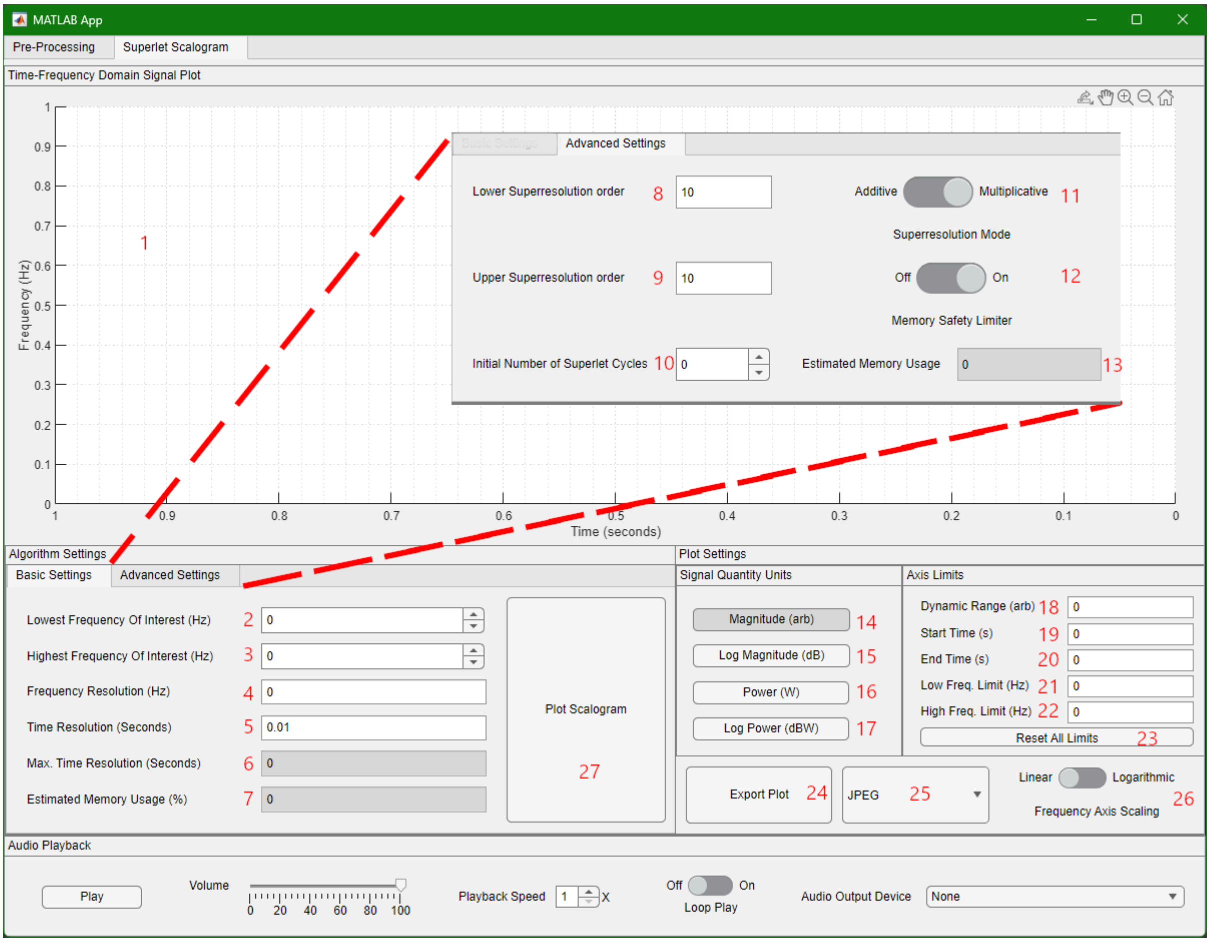
magnitude spectrum plot, so that the maximum amplitude and magnitude values are equal to 1. This is recommended when absolute source amplitude information is not required, or recording equipment sensitivity calibration values are not known. This does not affect the audio signal sent to the playback panel. Disabled after normalising, re-enabled by “Reset All”.

13. Re-scale the magnitude spectrum plot so that the frequency axis (y-axis) is logarithmically spaced. This has no effect on the SLT.
14. Switch the scaling of the magnitude spectrum y-axis units from linear magnitude to logarithmic magnitude (aka. Decibels/dB). This has no effect on the SLT.
15. Reset the audio signal start and end times, sample rate and amplitude values to their original values. This also clears any data in the SLT tab and returns settings there to default values.
16. Select the image file format for plot exports.
17. Export the time domain waveform plot and frequency domain magnitude spectrum plots as images.

4.3. The Audio Playback Panel

18. Play the audio signal. One click (button down) starts playback, click again (button up) to stop playback.
19. Set the playback volume. Changes to this setting do not affect the signal to be analysed by the SLT.
20. Set the playback speed. Increasing playback speed is common practice to make very low frequency signals easier to hear on conventional speakers. Changes to this setting do not affect the signal to be analysed by the SLT.
21. When this switch is set to “on”, playback of the signal will run in a continuous loop until the play button or the loop switch are set to “off”.
22. Select the audio output hardware device. Default will be the current Windows playback device.

[Contact the Developer: b.jancovich@unsw.edu.au](mailto:b.jancovich@unsw.edu.au)



4.4. The Superlet Scalogram Tab

1. Superlet Scalogram plot area. The results of the time-frequency analysis are visualised here.
2. The lowest frequency analysed by the SLT. Default is 10 Hz.
3. The highest frequency analysed by the SLT. Default is half of the sample rate.
4. The size of the frequency steps in the SLT analysis. A smaller number indicates higher resolution. Default is 0.25 Hz.
5. The size of the time steps in the SLT analysis. A smaller number indicates higher resolution. Default is the signal duration in seconds, divided by two times the sample rate.
6. The maximum possible time resolution (determined by the sample rate of the audio).
7. An estimate of the memory used to compute and render the visualisation.
8. The lower limit of superresolution orders. Recommend using default setting of “10”. (See below for further explanation)
9. The upper limit of superresolution orders. Recommend using default setting of “40”. (See below for further explanation)
10. Number of cycles in the base wavelet of a Superlet set. Recommend using default setting of “4”. (See below for further explanation)
11. Switch between additive or multiplicative superresolution. Recommend using default setting of “multiplicative”.
12. Memory Limiter Override Switch. Not recommended. May cause BASSA to crash or freeze.
13. Same as (7)
14. Set plot intensity units to unitless magnitude, normalized such that maximum dynamic range is 0 to 1.

Contact the Developer: b.jancovich@unsw.edu.au

15. Set plot intensity units to the logarithm of magnitude (Decibels/dB), normalized such that maximum dynamic range is -inf to 0.
16. Set plot intensity units to unitless Power, normalized such that maximum dynamic range is 0 to 1.
17. Set plot intensity units to the logarithm of Power (20*log10(magnitude)), normalized such that maximum dynamic range is -inf to 0.
18. Reduce the dynamic range of the SLT plot. This increases the lower axis limit of the colour bar, which represents intensity. It can help to de-emphasise low intensity sounds like background noise and can improve the visualisation of the signal of interest. Units follow the selections made in the "Signal Quantity Units" panel. E.g., reducing this value to 60dB with units set to mag dB will plot all intensities smaller than the maximum value minus 60dB, using the colour of the minimum value in the figure's colour axis.
19. Edit the time axis lower limit. This affects the plot only, and does not recalculate the SLT. The complete scalogram data remains intact when axis limits are reduced.
20. Edit the time axis upper limit. This affects the plot only, and does not recalculate the SLT. The complete scalogram data remains intact when axis limits are reduced.
21. Edit the frequency axis lower limit. This affects the plot only, and does not recalculate the SLT. The complete scalogram data remains intact when axis limits are reduced.
22. Edit the frequency axis upper limit. This affects the plot only, and does not recalculate the SLT. The complete scalogram data remains intact when axis limits are reduced.
23. Reset all plot axis limits.
24. Export the SLT Scalogram plot as an image.
25. Select the file type of the exported plot.
26. Re-scale the SLT plot so that the frequency axis (y-axis) is logarithmically spaced. This can improve visibility of closely spaced low frequency components.
27. Click to plot the scalogram using the current settings. Can be clicked again to re-plot with new settings.

Contact the Developer: b.jancovich@unsw.edu.au

5. Further Information

5.1. The Audio Playback Panel

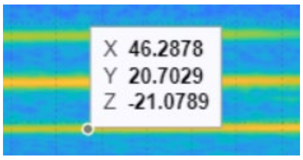
Any audio file that is loaded with a sampling rate that is higher or lower than the rates supported by the selected sound card are re-sampled to the closest supported sampling rate. This resampled version of the signal is used for audio playback only, and is not seen by the Superlet Scalogram tab.

The playback speed changes the sample rate of the audio output device. The file is not modified. The playback audio is always normalised such that the peak absolute amplitude is equal to 1. Volume adjustments are applied to normalised audio. This avoids the possibility of clipping the audio output.

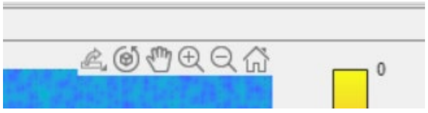
5.2. Interactive Plots

All plots in BASSA have interactive functionality. The responsiveness of these functions is highly dependent on memory usage.

Datatips: Clicking on a plotted line or surface sets a datatip, which will display values at those plot axis coordinates. For the waveform plot, X = time, Y = amplitude. For the magnitude spectrum plot, X = frequency, Y = magnitude. For the SLT plot, X = time, Y = frequency, Z = intensity. Right clicking a datatip opens a context menu with options for deleting current or all datatips.



Interactive Plot Controls: Hovering the mouse over the plot area will show several tool icons in the top right-hand side of the plot. These tools are, from left to right: Save/Export, rotate in 3D, Pan, Zoom in, Zoom out & Return to original view.



5.3. A Note on Performance and Memory Usage

BASSA was designed in the MATLAB App Designer, which uses the Chromium Embedded Framework (CEF) to render graphics. The CEF has a maximum JavaScript heap size of a little less than 4GB, and there is currently no way to increase the amount of memory it can use. This places an unfortunate limitation on BASSA's ability to plot very large SLT scalograms. Future MATLAB updates may resolve this, however at present, it is recommended that analysis settings be optimised for memory efficiency. Memory-saving default values have been set for the relevant controls. These can be overridden, but you may experience performance issues or instability.

Long audio files, files with a high sample rate, or analysing files across a large frequency range, or with very high time or frequency resolution will all cause memory use to skyrocket, and visualisations may take a very long time to compute and visualise. In extreme cases, this may cause BASSA to become unresponsive or crash entirely.

In testing, BASSA has been shown to operate normally with an audio file of 200 seconds in duration, with a sample rate of 500Hz, no trimming, no resampling, and all default SLT settings. The SLT in this test case took approximately 2 minutes to compute and plot on machine running Windows 11, with 64GB system memory, an Intel i7-1265U and an Nvidia T550 Laptop GPU.

Contact the Developer: b.jancovich@unsw.edu.au

To minimise memory usage issues, the following workflow is recommended:

1. Load an audio file.
2. Using the waveform plot as a guide, edit the start and end times and trim the audio, discarding region that contain no signals of interest.
3. Using the magnitude spectrum plot as a guide, determine the highest frequency that contains information of interest. Switching magnitude units to dB may be helpful here.
4. Reduce the sample rate to a value that is 2x the highest frequency of interest. E.g., if the highest information of interest in the audio signal occurs at 90 Hz, resample to 200 Hz, which will result in an upper frequency limit of 100 Hz.
5. Normalise the signal. Not a memory saver, just good practice.
6. Switch to the Superlet Scalogram Tab.
7. Inspect the memory estimate field.
8. If the memory estimate is above 3000 MB, it is recommended to reduce time or frequency resolution. While the default value for time resolution is calculated dynamically to optimise for memory efficiency, for many signals it may still end up being higher than necessary. Time resolution can often be reduced by an order of magnitude without losing meaningful temporal detail. For many animal sounds, a time resolution of 0.1 seconds is more than enough. Frequency resolution too, can often be reduced.
9. Use the edit fields for highest and lowest frequency of interest to further reduce the number of frequencies to be computed.
10. Recheck the memory estimate field. It will update when changes are made to frequency range and resolution settings.
11. Click “plot scalogram”.

5.4. Superlet Parameters: Some Technical Details

The Superlet transform is a generalisation of the continuous wavelet transform (CWT). The CWT determines the amount of energy at the frequency of interest is present in a signal, by convolving it with a “wavelet” whose centre frequency is at the frequency of interest. The initial “mother” wavelet has a specified number of cycles, and for each frequency of interest the mother wavelet is compressed in time, increasing its centre frequency while the number of cycles is held constant. The time resolution of the CWT is determined, in part, by the number of cycles in the mother wavelet.

The SLT operates by computing a plurality of continuous wavelet transforms, each having a different number of cycles in the mother wavelet. It then takes the geometric mean of the resulting CWT scalograms, to produce an SLT Scalogram.

The specific numbers of cycles used in the CWT set is determined by the frequency range, frequency resolution, initial number of Superlet cycles, the superresolution order interval, and whether the SLT is operating in additive or multiplicative mode.

Frequency Range & Resolution: The wider the frequency range and the higher the resolution (ie. the smaller and more numerous the frequency steps), the more CWTs are required to analyse the entire signal.

Initial Number of Superlet Cycles: The number of cycles in the mother wavelet of the initial CWT.

Multiplicative or Additive Mode: Determines whether the number of cycles for each CWT’s mother wavelet is calculated by using superresolution orders as multiplication factors or addition constants.

Contact the Developer: b.jancovich@unsw.edu.au

Upper and lower superresolution order: This defines the maximum and minimum number of cycles. If the frequency domain contains 0 Hz, then the lower order refers to frequency 0 and the upper order to the positive (upper) frequency boundary - the order will then mirror around 0 into the negative domain.

6. References

Moca, V. V., Bârzan, H., Nagy-Dăbâcan, A., & Mureșan, R. C. (2021). Time-frequency super-resolution with superlets. *Nature Communications*, *12*(1), Article 1. https://doi.org/10.1038/s41467-020-20539-9

Jancovich, B. (2023). A New Method for Visualising Animal Sounds. *Under Publication Review.*

Contact the Developer: b.jancovich@unsw.edu.au