Using a Sonic Time-Lapse Method as a Compositional Tool

FELIPE OTONDO and VICTOR POBLETE

Institute of Acoustics, Universidad Austral de Chile, Chile
Emails: felipe.otondo@uach.cl; vpoblete@uach.cl

This article describes an innovative compositional method based on the use of a sonic time-lapse algorithm to create soundscape audio montages. The method is based on the superposition of short audio samples of 24-hour continuous field recordings carried out in various kinds of wildlife sonic environments. The optimisation of the algorithm focused on the enhancement of gradual crossfade transitions between recorded samples and the use of variable sample durations as a way of recreating a natural sense of evolution of sonic events in time. Future developments of the project will integrate virtual reality interfaces and environmental education projects as part of the time-lapse algorithm machine-learning techniques and also creative tools suitable for multimedia installations.

1. INTRODUCTION

Recent developments in soundscape ecology research have made new creative tools and methodologies available that electroacoustic music composers and sound artists have successfully incorporated into their artistic practice (Carlyle 2007: 4–5; Demers 2010: 120–4; Lane and Carlyle 2013: 9–13; Bianchi and Manzo 2016: x–xxv). Field recordings have become an important source material for educators, programmers and artists devising electroacoustic compositions, sound installations, virtual reality applications and the sonification of raw data material (Westerkamp 2002: 51–2; Lane and Carlyle 2013: 9–13; Polli 2016: 3–7). Probably one of the most original features of these new tools is that they have been conceived as taking the active role of listeners as a point of departure in the creative process (Blesser and Blesser 2007: 176–8; Barclay 2017: 145–8). Inspired by these recent trends, regular periodic 24-hour continuous field recordings were carried out in different kinds of natural sonic environments aimed at exploring listener's perception of evolving sonic events over long periods of time (Otondo 2018a: 50). The results of these pilot tests inspired the idea of developing a sonic time-lapse method (STL) aimed at creating short audio montages to encapsulate specific acoustic events occurring in a particular location during a 24-hour cycle. The following article describes the design and optimisation of this method taking into consideration technical and artistic features of soundscape and electroacoustic composition research.

2. SONIC TIME-LAPSE

In recent decades sound artists and researchers from different backgrounds have developed projects involving the use of an STL method for creative and research purposes. The American composer and acoustic ecology researcher Bernie Krause implemented an acoustic monitoring method based on sound montages created using long continuous field recording extracts carried out in forests in California as a way of assessing the impact of acoustic pollution and deforestation (Krause 2004: 33–4). Like other acoustic ecology monitoring techniques, this method provides an effective tool to measure the impact of anthropogenic, biophonic and geophonic activity in natural habitats over long periods of time, but is a limited creative tool due to the low audio quality of the long continuous recordings obtained (Farina 2013: 12–3; Krause 2015: 33–4; Farina and Gage 2017: 8–9). Another artist who has explored the use of STL within a compositional context is Barry Truax. In collaboration with the members of the World Soundscape Project (WSP), Truax carried out short field recordings at various times in Vancouver Bay, aiming to capture subtle acoustic fluctuations throughout the day (Truax 2002: 7–9). The obtained recordings were then edited and assembled chronologically in an STL which captured timbral and spatial variations in the sonic landscape generated by changes in tides and movement of sea birds across the bay (Truax and Barret 2011: 1201–5; Truax 2012: 193–6). While this method involved an original montage approach and provided interesting sound material that was later used as the basis for soundscape compositions carried out by Truax and members of the WSP, it is constrained by the random character of weather conditions and the specific positions of the microphones in the bay. Years later the British sound recordist Chris Watson developed a compositional method based on the use of high-quality continuous stereo field recordings carried out over long periods of time in the Savannah in

Kenya (Watson 2003). Samples from these recordings were then analysed and manually assembled chronologically into a sound montage that aimed to encapsulate the temporal evolution of the main sonic events occurring over a 14-hour period. This time compression composition technique was the basis of the piece *Ol-olool-o* which is a carefully crafted abstract recreation of the temporal evolution of sonic events in a particular location where wild animals gather. This technique was later exemplified during a visit Watson made to Lancaster University in England where he collaborated with one of the authors and postgraduate students to carry out field recordings at various locations around the main campus (Otondo 2017: 93). The resulting audio material was later used as the basis for a multichannel sound installation presented at the university's concert hall with a very positive audience response (Watson 2016; Otondo 2018c: 136).

3. PILOT TESTS

Taking some of the examples and tests mentioned above as inspiration, it was decided to carry out various pilot tests aimed at developing an STL method that could be used for different kinds of creative purposes. In order to obtain sound material that could be used as the basis of the time-lapse process, a series of 24-hour continuous field recordings was carried out on the outskirts of three wetlands of the city of Valdivia in Chile, using a spaced-pair stereo recording setup with two omnidirectional microphones. By visually inspecting a spectrogram of the recorded material, twenty-four short 12-second audio samples were manually selected for each recorded hour focusing on the more prominent acoustic events taking place at each wetland. The audio samples obtained were then carefully edited by adjusting the dynamic envelope transitions between samples using 2-second linear crossfades in order to generate smooth and gradual aural transitions. The obtained samples were then assembled chronologically to produce a 4-minute time-lapse audio file aimed at recreating the temporal evolution of the main sonic events over the 24-hour recorded period (Otondo 2018b: 99). On the project's website visitors can listen to time-lapse audio files and samples of the 24-hour recordings of the three studied wetlands (Torres and Otondo 2019). A considerable resemblance can be observed when comparing the spectrograms of the original 24-hour recordings with the generated STL audio files. In this case the audio montage provides in a matter of a few minutes an acoustic summary of the main anthropogenic, biophonic and geophonic activity in each wetland during a 24-hour cycle. What appears to be a static soundscape with sparse acoustic activity in the original field recordings becomes dynamic and fresh when presented in a compressed temporal framework. While this manual method provides a straightforward way of interacting with recorded sound material through careful editing and selection, it has various practical limitations when implemented in a professional creative environment. The large size of the original 24-hour continuous field recordings audio files makes editing and manipulation of sound material lengthy and highly dependent on the computer power available. Another critical point is the curve shape of crossfade dynamic transitions between audio samples for each recorded hour (Roads 1996: 769). The use of default linear crossfade transitions between audio samples create a decrease in the added overall RMS intensity levels manifesting in regular dynamic variations (Roads 2001: 200). These periodic intensity variations become ostensibly audible in the finalised time-lapse audio montage, especially in terms of background noise and high frequency fluctuations. Taking into account some of these limitations, an automated computer algorithm that could incorporate and optimise the STL methodology described previously was designed and implemented.

4. SONIC TIME-LAPSE ALGORITHM DESIGN AND OPTIMISATION

The STL algorithm was planned using a set of computing techniques conceived to improve the reliability of the montage method. The design focused on achieving an automated chronological superposition of short audio files, improving crossfade transitions between recording samples and the implementation of a graphical visualisation tool to explore raw input data. It includes a modularised code which allows document design, simultaneous assessment of the script by several programmers and bug detection and can easily adapt to specific user requests (Wilson et al. 2014). The code is written in Python 3.7.4 and integrated into a Jupyter notebook web application to allow for the underlying theory and equation information to be integrated as part of the STL implementation (Kluyver et al. 2016). This integrated environment allows the code to be separated into cell blocks that can be recomputed dynamically, allowing the user to work on specific points at a time without the need to recompute the code written before that point. The STL code also incorporates scientific open-source libraries such as Numpy (van der Walt, Colbert and Varoquaux 2011); Pandas (McKinney 2010); Scipy (Jones et al. 2001); Matplotlib (Hunter 2007); SoundFile (PySoundFile 2019) and Tkinter (Python Tkinter 2019). In order to deal with the visualisation of a considerable volume of sonic samples generated by the processing of large recordings,
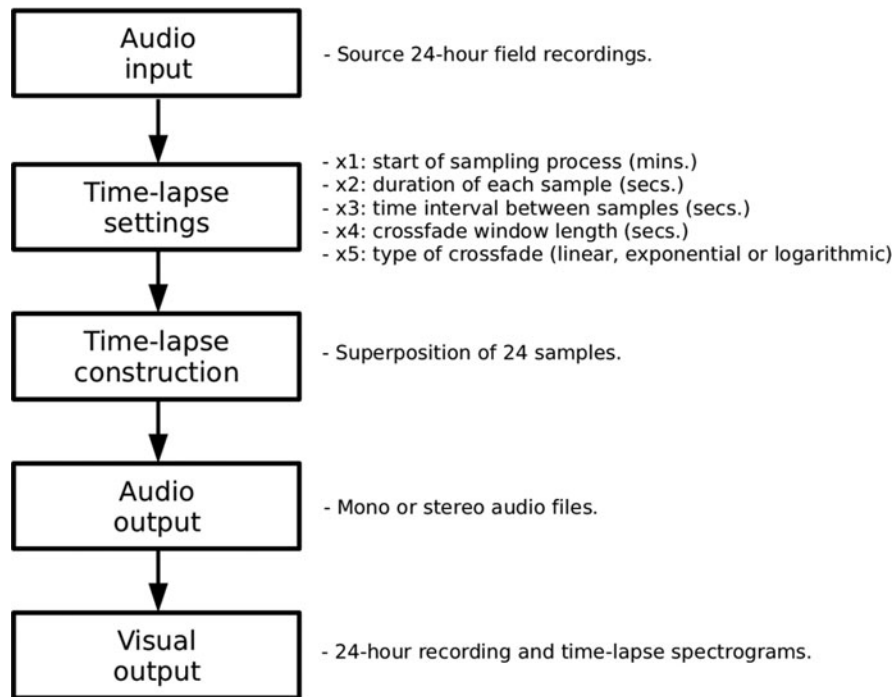
Figure 1. Sonic time-lapse algorithm overview.
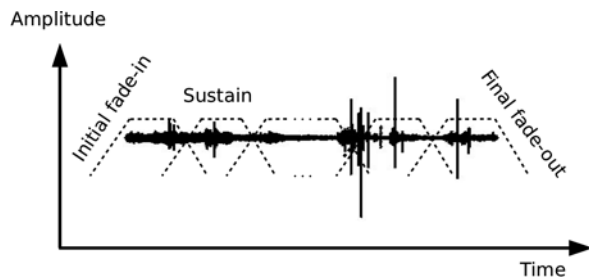
Figure 2. Time-lapse crossfade montage process.

spectral visualisation tools were designed using Bokeh, Holoviews and Datashader plotting libraries. Figure 1 illustrates a diagram of the STL algorithm and Figure 2 shows dynamic envelopes of the crossfading process between consecutive recorded samples. The input field recording in this case is a stereo audio wav file sampled at 44.1 kHz and quantised with 16 bits. The first input parameter is the minute of each hour where the sampling process takes place. As shown in the figure, the second parameter is the duration of each sampled segment. The third parameter is the time interval between samples and the fourth is the crossfade window duration between two consecutive segments. The last parameter allows control over the type of crossfade used between recorded samples: linear, exponential or logarithmic. As mentioned previously, the goal in this case was to provide subtle

crossfade transitions that maintain the same overall RMS level in order to avoid noticeable audible dynamic fluctuations between segments. Once these parameters have been input, the algorithm builds the STL by concatenating in time the 24 audio samples obtained using the montage method described in the previous section. The final output of the algorithm is a mono or stereo audio file labelled according to crossfade and sustain durations. In this case a file with a 4-second crossfade and a 6-second sustain duration would be represented as '4–6–4'. The last stage of the algorithm routine allows users to visualise the spectrogram of the generated file and listen to the outcome.

5. CREATIVE APPLICATIONS

Practical tests with the algorithm were carried out using as source material 24-hour field recordings that took place at the Parque Urbano El Bosque wetland in March 2019, in the midst of frog mating season. Following a similar approach to the manual sampling process carried out for the pilot tests described previously, short samples of every recorded hour were blended chronologically to generate time-lapse audio files of different durations. Sustain and crossfade durations of samples were modified as a way of assessing the perceptual and artistic impact of various temporal frameworks. Table 1 shows crossfade, sustain and overall durations for the four compared generated audio files and Figure 3 shows spectrograms of the

Table 1. Temporal features of generated time-lapse audio files.

| Time-lapse example	Crossfade duration	Sustain duration	Overall montage duration
A	2″	2″	46″
B	3″	5″	1′31″
C	6″	10″	3′02″
D	12″	20″	6′04″

generated files compared with the 24-hour recordings (bottom figure). Time-lapse audio files of spectrograms A, B, C and D shown in the figure can be heard in Sounds examples 1, 2, 3 and 4, respectively. Informal listening comparisons were carried out by equating the generated audio files. The shorter the sustain and crossfade durations, the more abstract and less recognisable the sound sources become in the final time-lapse montage. In this case dynamic, timbral and spatial articulations are perceived as very intense and regular due to the fact that the attacks of the recorded sonic events are mostly captured in the short sampling window for each hour making the final audio montage less natural and slightly monotonous in terms of frequency of occurring events. As crossfade and sustain durations increase, the sharp nature of attacks of sonic events becomes less frequent and nuanced. The succession of sonic events becomes less regular and therefore a greater variety of timbres and spatial depth starts to emerge. Overall dynamic articulations become varied and less extreme, providing a more subtle and sparse sonic character to the resulting montage. As durations become even longer, the generated montage starts to resemble portions of the original 24-hour recordings in terms of timbral and spatial qualities. Sounds of birds and barking dogs that were unrecognisable in the shorter audio montage now become clearly recognised. These early results suggested that the algorithm could be developed further as a standalone creative application that could be tailored for different kinds of creative purposes. In a similar fashion to granular synthesis and convolution methods being applied to soundscape composition, the time-lapse application could be adapted to emphasise macro- or micro-temporal features of long continuous field recordings (Kendall 2010: 68–71). In some cases users might be interested in emphasising easily recognisable sound sources occurring over large periods of time while in other cases they might want to highlight timbral features of short sound impulses (Truax 2012: 197–8). By simply selecting the overall duration of the finalised audio files, users can create an audio montage that could emphasise timbral aspects and be used as short sequences of abstract intense sounds or longer

soundscape files with a more gradual and natural sense of development. This could provide an interesting alternative to the use of time-stretching techniques for soundscape composition which in many cases tends to distort original recordings by adding undesired artefacts to the audio signals due to the complex acoustic features of field recordings. By organically altering temporal features of sampled sounds such as crossfade and sustain durations, users could create short or long sound files that could encapsulate the more important acoustic attributes of long field recordings. Machine-learning techniques used in the prediction of spatial and temporal features of soundscapes could also be incorporated into the algorithm script (Farina and Gage 2017: 217; Bellisario and Pijanowski 2019). The integration of neural networks as part of the model could provide a suitable training method to avoid the arbitrary selection of samples for each recorded hour. This could allow the possibility of training the STL model according to a set of acoustic and perceptual features that could shape the final audio montage for artistic, eco-acoustics or soundscape quality assessment purposes (Boes, Filipan, De Conensel and Botteldooren 2017: 122–5). In line with new ISO standards about soundscape applications the use of Ambisonics and binaural recording techniques could also be used to explore three-dimensional qualities of immersive natural and urban acoustic environments (Toole 2008: 285–6; ISO 12913–1 2014: 1, ISO 129131–2 2017: 25–7). The goal in this case would be to investigate the recording/reproduction limitations and advantages of both techniques in relation to creative applications based on the STL method. In this case immersive multimedia site-specific installations for art galleries, concerts or museums would be the obvious choice, but more innovative applications such as a virtual reality software for visually impaired people or augmented reality game audio tools could probably open new perspectives for the development of the STL method.

6. CONCLUSION AND FUTURE WORK

The proposed sonic time-lapse method is an innovative creative tool aimed at opening new possibilities for composers and sound artists interested in large-scale transformations in the time domain. In line with research that explores the role of soundwalking as a listening practice and soundscape composition as a way of evoking memory, the STL method presented here provides an innovative creative tool for exploring perceived spatio-temporal changes of acoustic environments (Kolber 2002: 41–3; ISO 129131–2 2017: 13–24; Martin 2018: 21–7). The extraction of salient sonic events of a 24-hour continuous field recording
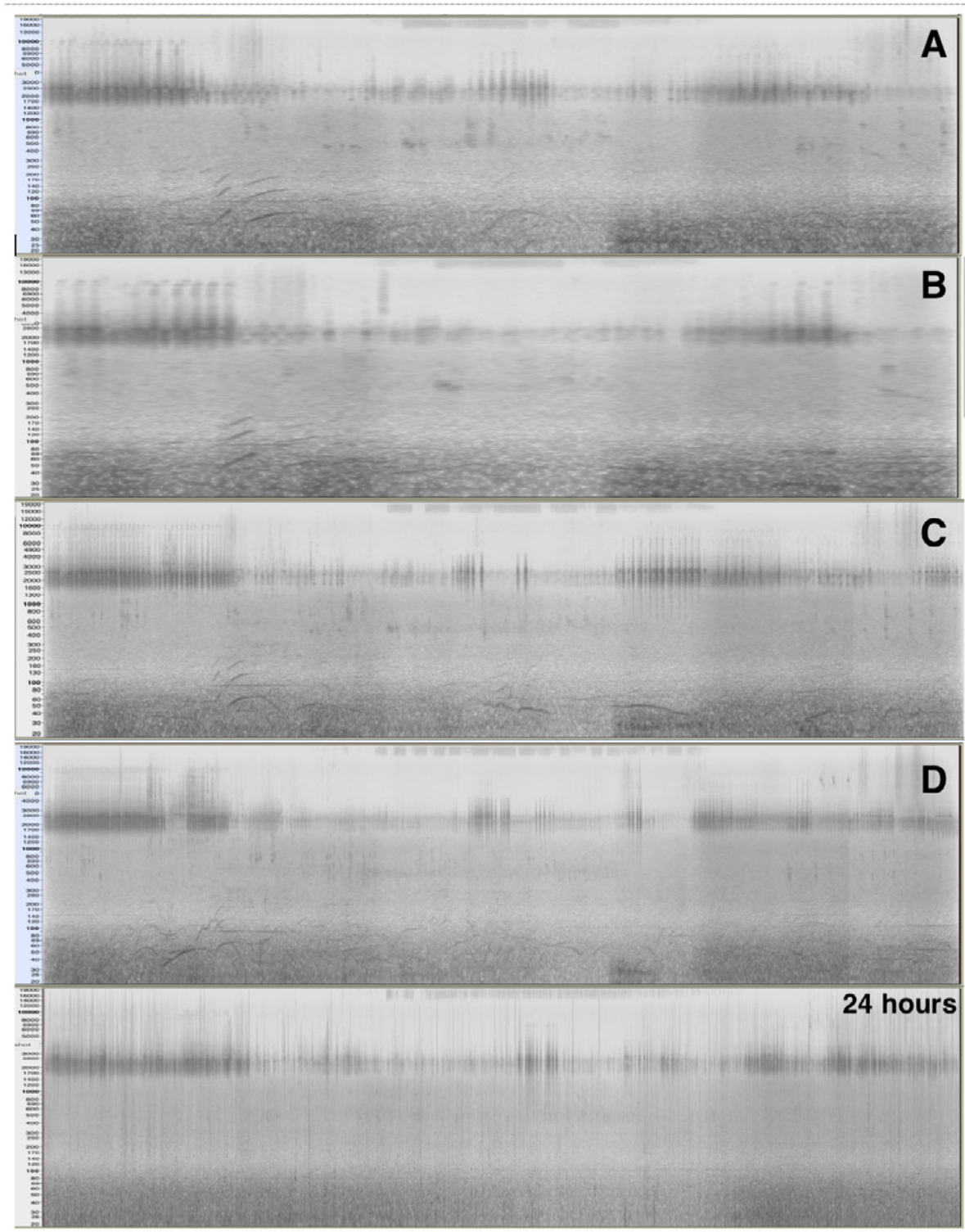
Figure 3. Spectrograms of generated time-lapse audio files and original 24-hour field recording (bottom).

in a short audio montage allows the possibility of creating original sonic constructions where dynamic, timbral and spatial features of a particular soundscape can be reinforced or freely transformed. The proposed STL algorithm discussed in this article shows clear improvements from early manual versions in terms of the stabilisation of dynamic fluctuations of crossfade transitions. Future developments of this project will consider improvements to the time-lapse algorithm by incorporating a machine-learning

application to adjust the sampling process to specific types of acoustic events recorded. Eco-acoustics signal processing techniques could be considered as training tools for the algorithm in order to select and tag the most relevant sonic events in the recordings to be included in the final audio montage. This could open new alternatives for creating audio montage files with less arbitrary presets and the possibility of a stand-alone application where users could design their own time-lapse montage files according to a set of variable and fixed temporal parameters. The use of audio signals in various spatial formats as source material for the algorithm will also be considered. The 24-hour field recordings in Ambisonics and binaural audio formats will be used to devise and implement immersive sonic environments in the form of multichannel installations. A virtual reality educational tool will also be considered as a way of providing children with limited mobility the possibility of familiarising remotely with wetland natural environments.

Acknowledgements

SUPPLEMENTARY MATERIAL

To view supplementary material for this article, please visit https://doi.org/10.1017/S1355771820000102

REFERENCES

Barclay, L. 2017. Listening to Communities and Environments. *Contemporary Music Review* **36**(3): 143–58.

Bellisario, K. M. and Pijanowski, B.C. 2019. Contributions of MIR to Soundscape Ecology. Part I: Potential Methodological Synergies. *Ecological* **51**: 96–102.

Bianchi, F. and Manzo, V. J. 2016. *Environmental Sound Artists: In Their Own Words.* New York: Oxford University Press.

Blesser, B. and Blesser, R. 2007. *Spaces Speak, Are You Listening?* Cambridge, MA: MIT Press.

Boes, M., Filipan, K., De Coensel, B. and Botteldooren, D. 2017. Machine Listening for Park Soundscape Quality Assessment. *Acta Acustica united with Acustica* **104**(1): 121–30.

Carlyle, A. 2007. *Autumn Leaves: Sound and the Environment in Artistic Practice.* Paris: Double Entendre.

Demers, J. 2010. *Listening Through the Noise: The Aesthetics of Experimental Music.* New York: Oxford University Press.

Farina, A. 2013. *Acoustic Ecology.* Dordrecht: Verlag.

Farina, A. and Gage, S. 2017. *Ecoacoustics.* Oxford: Wiley.

Hunter, J. D. 2007. Matplotlib: A 2D graphics environment. *Computing in Science & Engineering* **9**(3): 90–5.

ISO 12913–1. 2014. *Acoustics – Soundscape – Part 1: Definition and Conceptual Framework.* Geneve, Switzerland: International Organization for Standarization.

ISO 12913–2. 2017. *Acoustics – Soundscape – Part 2: Data Collection and Reporting Requirements.* Geneve, Switzerland: International Organization for Standarization.

Jones, E., Oliphant, T. and Peterson, P., 2001. SciPy: Open source scientific tools for Python. www.scipy.org/ (accessed 6 September 2019).

Kendall, G. 2010. Meaning in Electroacoustic Music and the Everyday Mind. *Organised Sound* **15**: 63–74.

Kluyver, T., Ragan-Kelley, B., Pérez, F., Granger, B., Bussonnier, M., Frederic, J., et al. 2016. Jupyter Notebooks – A Publishing Format for Reproducible Computational Workflows. In F. Loizides and B. Schmidt (eds.) *Positioning and Power in Academic Publishing: Players, Agents and Agendas.* Amsterdam: IOS Press, 87–90.

Kolber, D. 2002. Hildegard Westerkamp's Kits Beach Soundwalk: Shifting Perspectives in Real World Music. *Organised Sound* **7**(1): 41–3.

Krause, B. 2004. *Wild Soundscapes: Discovering the Voice of the Natural World.* Berkeley: Wilderness Press.

Krause, B. 2015. *Voices of the Wild.* London: Yale University Press.

Lane, C. and Carlyle, A. 2013. *In the Field: The Art of Field Recording.* Devon: Uniformbooks.

Martin, B. 2018. Soundscape Composition: Enhancing Our Understanding of Changing Soundscapes. *Organised Sound* **23**(1): 20–8.

McKinney, W. 2010. Data Structures for Statistical Computing in Python. *Proceedings of the 9th Python in Science Conference.* Austin, Texas, 51–6.

Otondo, F. 2017. Context-based Composition in an Interdisciplinary Collaborative Framework. *Organised Sound* **22**(1): 93–100.

Otondo, F. 2018a. Listening to Wetland Soundscapes. *Leonardo Music Journal* **28**: 50–2.

Otondo, F. 2018b. Valdivia's Wetland Soundscape. *Leonardo Music Journal* **28**: 99.

Otondo, F. 2018c. Paisajes sonoros reales e imaginarios. *Resonancias* **22**(42): 131–41.

Polli, A. 2016. Sonification of Global Environmental Data. In F. Y. Bianchi and V. J. Manzo (eds.) *Environmental Sound Artists: In Their Own Words.* New York: Oxford University Press, 3–8.

PySoundFile. 2019. https://pypi.org/project/PySoundFile (accessed 6 September 2019).

Python Tkinter. 2019. https://docs.python.org/3/library/tkinter.html. (accessed 6 September 2019).

Roads, C. 1996. *The Computer Music Tutorial.* Cambridge, MA: MIT Press.

Roads, C. 2001. *Microsound.* Cambridge, MA: MIT Press.

Toole, F. 2008. *Sound Reproduction.* Oxford: Focal Press.

Torres, R. and Otondo, F. 2019. *Soundlapse Proyect Website.* www.soundlapse.net (accessed 6 September 2019).

Truax, B. 2002. Genres and Techniques of Soundscape Composition as Developed at Simon Fraser University. *Organised Sound* **7**(1): 5–14.

Truax, B. 2012. Sound Listening and Place: The Aesthetic Dilemma. *Organised Sound* **17**(3): 193–201.

Truax, B. and Barrett, G. 2011. Soundscape in a Context of Acoustic and Landscape Ecology. *Landscape Ecology* **26**(9): 1201–7.

van der Walt, S., Colbert, S. C. and Varoquaux, G. 2011. The Numpy Array: A Structure for Efficient Numerical Computation. *Computing in Science & Engineering* **13**(2): 22–30.

Watson, C. 2016. *Chris Watson's website*. http://chriswatson.net/2011/10/14/a-journey-south-9th-november-2011/.

Westerkamp, H. 2002. Linking Soundscape Composition and Acoustic Ecology. *Organised Sound* **7**(1): 51–6.

Wilson, G., Aruliah, D. A., Brown, C. T., Chue Hong, N. P., Davis, M., Guy, R. T., et al. 2014. Best Practices for Scientific Computing. *PLoS Biol* **12**(1): e1001745.

DISCOGRAPHY

Watson, C. 2003. *Weather Report*. London: Touch Music.