Computing biodiversity change via a soundscape monitoring network

Tzu-Hao Lin[*], Yu Tsao[§]
Academia Sinica
Taipei, Taiwan (R.O.C.)
[*]schonkopf@gmail.com
[§]yu.tsao@citi.sinica.edu.tw

Yu-Huang Wang
Taiwan Academy of Ecology,
Changhua, Taiwan (R.O.C.)

Han-Wei Yen
Academia Sinica Grid
Computing Centre
Taipei, Taiwan (R.O.C.)

Sheng-Shan Lu
Taiwan Forestry Research
Institute,
Taipei, Taiwan (R.O.C.)

***Abstract*—A monitoring network for biodiversity change is essential for wildlife conservation. In recent years, many soundscape monitoring projects have been carried out to investigate the diversity of vocalizing animals. However, the acoustic-based biodiversity assessment remains challenging due to the lack of sufficient recognition database and the inability to disentangle mixed sound sources. Since 2014, an Asian Soundscape monitoring project has been initiated in Taiwan. So far, there are 15 recording sites in Taiwan and three sites in Southeast Asia, with more than 20,000 hours of recordings archived in the Asian Soundscape. In this study, we employed the visualization of long-duration recordings, blind source separation, and clustering techniques, to investigate the spatio-temporal variations of forest biodiversity in the Triangle Mountain, Lienhuachih, and Taipingshan. On the basis of blind source separation, biological sounds, with prominent diurnal occurrence pattern, can be separated from the environmental sounds without any recognition database. Thus, clusters of biological sounds can be effectively identified and employed to measure the daily change in bioacoustic diversity. Our results show that the bioacoustic diversity was higher in the evergreen broad-leaved forest. However, the seasonal variation in bioacoustic diversity was most evident in the high elevation coniferous forest. This study demonstrates that a suitable integration of machine learning and ecoacoustics can facilitate the evaluation of biodiversity changes. In addition to biological activities, we can also measure the environmental variability from soundscape information. In the future, the Asian Soundscape will not only serve as an open database for soundscape recordings, but also will provide tools for analyzing the interactions between biodiversity, environment, and human activities.**

Keywords—Biodiversity assessment; soundscape; biological sounds; machine learning; open data

I. INTRODUCTION

Information on biodiversity change is essential for conservation and wildlife management. However, field surveys of biodiversity are time-consuming, requiring a significant amount of resources. In recent decades, numerous automatic sensing techniques have been developed to facilitate field observation in remote areas [1]. Satellite remote sensing has been employed to investigate changes in land cover and use over time [2]. Camera traps have also been applied to investigate the behavioral pattern of wildlife in remote areas [3]. Although these sensing techniques have considerably reduced the difficulty in studying large-scale assemblage-environment relationships, it continues to be difficult to investigate the changes in animal community on a daily basis.

Passive acoustic monitoring has been widely employed to study the occurrence and behavior of vocalizing animals, such as birds, bats, fish, and whales [4]–[7]. By deploying an autonomous sound recorder, it is possible to record various animal vocalizations, investigating the dynamics of species diversity in remote forests or deep waters. In addition to biological sounds, a sound recorder can also collect environmental and anthropogenic sounds, which are the essential components for the soundscape [8], [9]. Therefore, the soundscape can be considered as an acoustic sensing information for us to explore and investigate the interactions among the habitat environment, biodiversity, and human development.

Many soundscape monitoring networks have been established in recent years, such as Soundscape Ecology Research Projects (http://centerforglobalsoundscapes.org/pumilio-sites) [8], Cyberforest (http://global.pioneer/en/crdl_design/soundlab/cyberforest) [10], and Listening to the deep ocean (http://listentothedeep.com/acoustics) [11]. Although a soundscape monitoring network appears to be a promising solution for biodiversity assessment, the data analysis remains challenging because all the data collected by autonomous recorders are without corresponding labels. More specifically, there is no information available on the species of vocalizing animals corresponding to the sounds in a recording for analyzing biodiversity. In practice, it is nearly impossible to manually label all vocalizing animals from the enormous amount of field recordings. To facilitate data analysis, many researchers have developed rule-based detectors and supervised classification algorithms to help identify target species [12]–[15]. However, the automatic detection and classification performances highly depend on the signal-to-noise ratio of field recordings and the amount of training data [16]. At present, a comprehensive acoustic library for identifying wildlife is not available. Therefore, previous bioacoustics studies can only focus on a limited number of target species, instead of the entire community of vocalizing animals.

To study the entire community of vocalizing animals, researchers have developed several ecoacoustics indices to quantify field recordings. Although each ecoacoustics index

---

[*] Supported by Independent Postdoctoral Research Fellowship, Ministry of Science and Technology, Taiwan (R.O. C.)

has its unique formulation, all the indices share the same concept: evaluating the variation in biological activities by measuring the spectral and temporal complexities [17]–[19]. However, field recordings are influenced by simultaneous sound sources, e.g., weather noise, traffic noise, and construction noise. Without proper denoising, the acoustic-based biodiversity assessment may be biased by the presence of environmental and anthropogenic sounds.

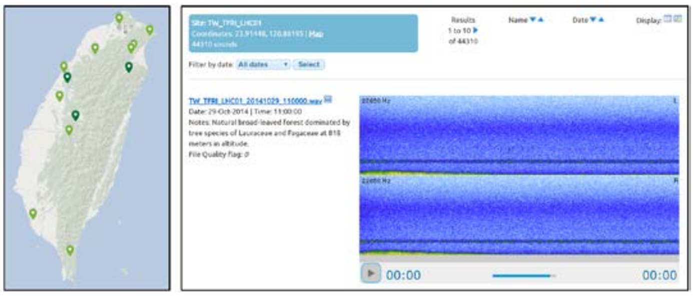
In this study, we sought to address the potential application of computing biodiversity change by using long-duration field recordings. To establish a long-term soundscape monitoring network, challenges such as the establishment of recording sites, management of big data, and the analysis of soundscape information must be overcome. Therefore, we (1) outlined the architecture of the Asian Soundscape with its autonomous recording systems and open data platform, (2) demonstrated the analysis of bioacoustic diversity in three forests at different altitudes, (3) discussed the significance and challenges associated with using soundscape information in biodiversity monitoring.

II. METHODS

A. *Asian Soundscape Monitoring Network*

In Taiwan, a soundscape monitoring network has been established by the Taiwan Forestry Research Institute and the Research Center for Information Technology Innovation, Academia Sinica. Long-duration field recordings were acquired by using Song Meter SM2+ and SM4 acoustic recorders (Wildlife Acoustics, Inc., Maynard, MA). Song Meter recorders are compact weatherproof recorders, equipped with two microphones. The built-in microphones for SM2+ and SM4 are slightly different in sensitivity (−33.5 and −36 dB re 1 V/Pa for SM2+ and SM4, respectively). However, both of them perform well in the frequency range of 100 Hz to 10 kHz. We deployed SM2+ and SM4 recorders in urban parks, terraces, wetlands, and forests to monitor the soundscape in different habitat environments (Fig. 1). Each recorder was scheduled to run a 5-minute recording, with 44.1 kHz sampling frequency in two channels, every half an hour. To ensure data recovery, we changed memory cards, batteries, windscreens, and broken microphones every 2.5 months for SM4 recorders. For SM2+ recorders, we had to service the recorders every month due to the higher power consumption.

According to the schedule, the data volume of each site increases by approximately 1 TB every year. To support the management of acoustic data, the Academia Sinica Grid Computing Centre has established the online platform of Asian Soundscape (http://soundscape.twgrid.org). On the basis of the distributed computing infrastructure, we implemented the Pumilio application for managing and visualizing acoustic recordings [20]. On the Asian Soundscape, users can listen to all the archived audio clips, while investigating the spectral features of signals of their interest on spectrograms. Metadata, such as the latitude, longitude, habitat characteristics, recording date, recording time, and the sampling frequency of each audio clip have been saved for further analysis. Currently, there are 15 recording sites in Taiwan, with more than 20,000 hours of recordings, while the data volume continues to increase. In addition, the Taiwan Forestry Research Institute also collaborated with researchers in Malaysia, Thailand, and Vietnam to establish three recording sites in tropical forests. Through international collaboration, the Asian Soundscape serves as the first open soundscape database in the Southeast Asia.



TW_TFRI_LHC01_20141029_110000.wav
Date: 29-Oct-2014 | Time: 11:00:00
Notes: Natural broad-leaved forest dominated by tree species of Lauraceae and Fagaceae at 818 meters in altitude.
File Quality flag: 0

Fig. 1. Overview of the Asian Soundscape. The left panel shows the map of recording sites in Taiwan, while the dark green markers represent the three recording sites used in this study. The right panel shows the user interface and spectrograms of the left and right channels for an audio clip.

B. *Analysis of Soundscape Information*

To facilitate the acoustic-based biodiversity assessment, we employed three tools for the analysis of soundscape: (1) visualization of long-duration recordings using a long-term spectrogram (LTS), (2) unsupervised separation of soundscape components, and (3) modeling of bioacoustic diversity.

To visualize long-duration recordings, we measured the median and mean power spectrums of each 5-minute audio clip, and sequentially combining all the power spectrums of a recording site [21]. By using an LTS, biological sounds with evident diurnal occurrence can be efficiently visualized [22]. Some animals, such as cicadas and crickets, chorus for several hours. These choruses can be easily identified on a median-based LTS. On the other hand, most birds and mammals only produce transient calls, which cannot be detected on a median-based LTS. To visualize these transient calls, we calculated the difference between mean-based and median-based LTSs (hereafter difference-based LTS). Thus, we can investigate biological choruses and transient calls by using median-based LTS and difference-based LTS, respectively.

In addition to visualization, an LTS can also serve as the input data for separating different soundscape components. To separate soundscape components without any training data, we developed the periodicity-coded nonnegative matrix factorization (PC-NMF) algorithm to separate biological chorus and the other noise sources [23]. PC-NMF is a self-learning algorithm, which consists of two stages of NMF. In the first stage, the input spectrogram is factorized as a basis matrix and an encoding matrix. The basis matrix represents a collection of spectral components, while the encoding matrix describes the time activation of each basis. In the second stage, the diurnal periodicity of each basis is measured by conducting discrete Fourier transform on the encoding information. Two basis groups are clustered according to the periodicity by using a secondary NMF process. Finally, the biological chorus can be reconstructed by the spectral bases with a strong diurnal periodicity and the associated encoding information. Meanwhile, the noise components, such as weather noise and anthropogenic noise, can be suppressed without any recognition database.

After removing the unwanted noise from median-based LTS and difference-based LTS, different groups of biological sounds, with unique spectral features, were identified by using clustering techniques. Firstly, we used the principle component analysis to reduce feature dimension. Principle components, which pertain to more than 85% of the variance, were used as input features for the *k*-means clustering. Subsequently, we iteratively attempted different cluster numbers (*k*), choosing the clustering result when the mean intracluster variation (sum of point-to-centroid distances within each cluster) was reduced to 2.5% of the variation of the entire data set. The clustering procedures were applied to median-based LTS and difference-based LTS to investigate the structure of biological chorus and transient calls, respectively.

Finally, the modeling of bioacoustic diversity was conducted based on the clustering result of the median-based LTS and difference-based LTS. Among the various biological sounds, transient calls and choruses are not mutually exclusive. Even under the same chorusing conditions, different types of transient calls may still be detected. We numerically labeled each group of chorus and transient call. Then, each 5-minute recording clip was given a bioacoustic code by combining the numerical labels of chorus and transient calls. Thus, we could investigate the change in bioacoustic diversity by measuring the Shannon entropy, based on the probabilistic distribution of bioacoustic codes. The modeling of bioacoustic diversity was achieved by using a spline on the change in bioacoustic diversity in each month.

C. Computing Forest Biodiversity at Different Altitudes

Since October 2014, the Taiwan Forestry Research Institute had established three recording sites at Triangle Mountain (hereafter MLSY), Lienhuachih Research Center (hereafter LHC), and Taipingshan (hereafter YLTPS). The altitudes of the three sites are approximately 500, 800, and 1900 m, respectively. The recording site at MLSY is a subtropical, evergreen, broad-leaved forest, located in Sanyi Township, Miaoli County. This site is close to National Highway No. 1 (approximately 1.1 km), and the recorder was deployed next to a hiking trail. The forest environment of the second recording site is similar to that of the first. This site is located in Lienhuachih Research Center, Yuchi Township, Nantou County. Unlike MLSY, LHC can only be visited by researchers. Therefore, the level of human disturbance is significantly lower at LHC, compared to the other two sites. The recording site at YLTPS is a coniferous forest, located at a higher elevation in Taipingshan National Forest Recreation Area, Yilan County. The recorder was deployed next to a trail. In this study, we employed three analysis tools to investigate the spatio-temporal variations of soundscape and bioacoustic diversity in the three recording sites by using long-duration recordings from October 2014 to December 2016. All audio clips during this period are available on the website of Asian Soundscape.

III. RESULTS

Through the median-based LTSs, the spatial and temporal change of forest soundscape can be effectively compared (Fig. 2). At MLSY, low-frequency noise (< 2 kHz) was prominent during the entire recording period due to the proximity to National Highway No. 1. At MLSY and LHC, the soundscapes were more complex from April to November, based on the overall increase in power spectral densities. However, the forest soundscape was completely different at YLTPS. Similar seasonal change pattern was not observed in the median-based LTS at YLTPS, with the soundscape being dominated by broadband noise. Upon listening to the original audio clips, these broadband noises were determined to be caused by heavy rainfall. The influence of rainfall noise can also be observed on the difference-based LTS at YLTPS, with most of the transient calls being masked by the rainfall noise.

By using the PC-NMF, we separated biological choruses with strong diurnal occurrence and the other sound sources displayed in the median-based LTSs (Fig. 3). The results showed that biological choruses dominated the soundscape at MLSY and LHC. Moreover, spectral characteristics also varied among different seasons. On the contrary, except for the high-frequency components (> 14 kHz), minimal biological chorus could be recorded at YLTPS. The PC-NMF performed well to reduce the rainfall noise displayed on the difference-based LTSs at MLSY and LHC. It was not effective for the rainfall noise displayed on the difference-based LTS at YLTPS. Although the low-frequency component of rainfall noise had been removed, some noise components higher than 2 kHz continued to be visible. However, the PC-NMF provides us an efficient solution in enhancing biological sounds without any

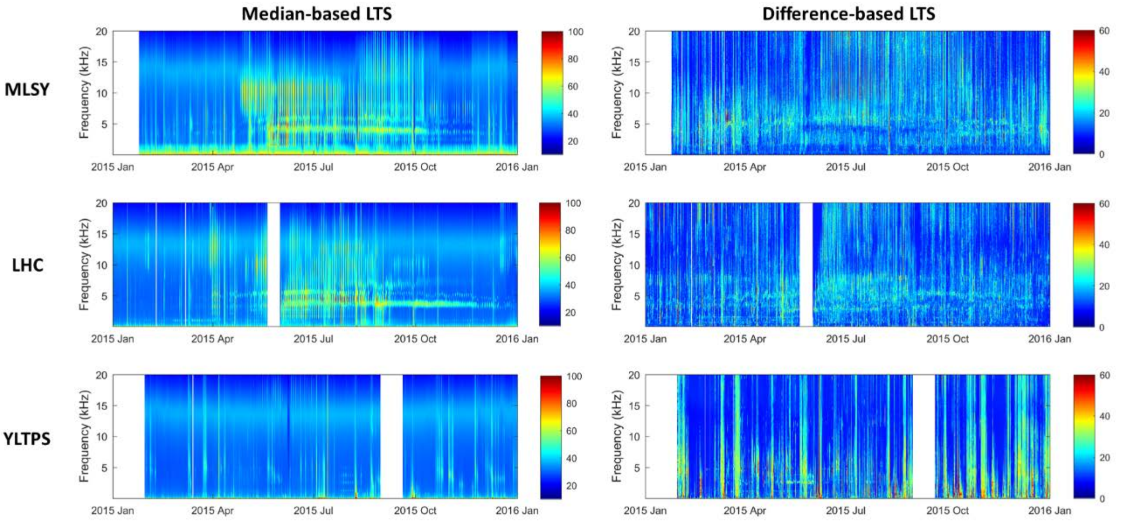


Fig. 2. Examples of median-based LTS (left panels) and difference-based LTS (right panels) at MLSY (upper panels), LHC (middle panels), and YLTPS (lower panels). Blank areas represent the period without recordings.

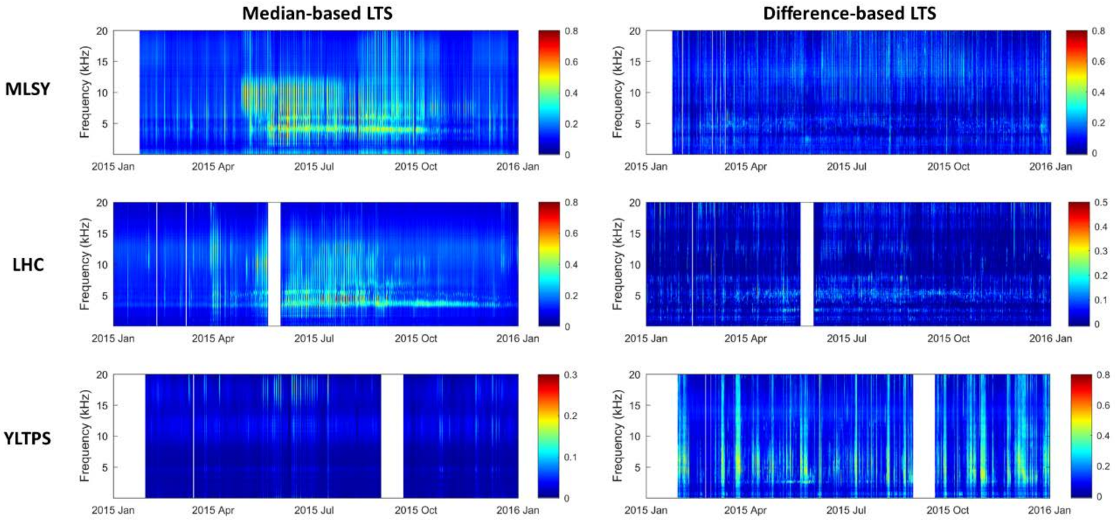


Fig. 3. Examples of using the PC-NMF to separate biological sounds from median-based LTS (left panels) and difference-based LTS (right panels) at MLSY (upper panels), LHC (middle panels), and YLTPS (lower panels). Blank areas represent the period without recordings. Note that some high-frequency components of rainfall noise displayed on the difference-based LTS at YLTPS were not removed by the PC-NMF.

sound recognition database.

In this study, the bioacoustic diversity was measured based on the clustering result of biological sounds. Without the PC-NMF, the clustering result was highly influenced by weather noise. Clusters associated with rainfall noise were commonly identified during winter at MLSY and LHC, as opposed to the entire year at YLTPS (Fig. 4). These clusters usually occurred in a pattern, lasting for several hours or several days, thus deviating the measurement of bioacoustic diversity. After the separation procedure, the influence of weather noise in the clustering result was not very prominent. In particular, for the data collected from YLTPS, the clustering result varied significantly before and after the separation procedure (Fig. 5).

The modeling result shows that the bioacoustic diversities were higher in the two evergreen broad-leaved forests than in the coniferous forest at a higher elevation. Moreover, the seasonal changing patterns of bioacoustic diversity also varied among different forests. At YLTPS, bioacoustic diversities increased from April to July. At LHC, bioacoustic diversities increased from April to October. However, bioacoustic diversities decreased only slightly from December to March at MLSY.

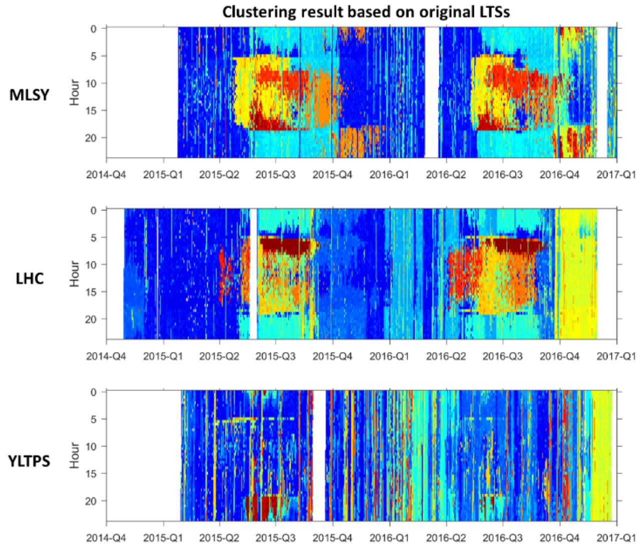


Fig. 4. Analysis results based on the original median-based LTS and difference-based LTS. Each color represents a different soundscape scene, with unique spectral features recognized by using *k*-means clustering. Without appropriate denoising, the influence of rainfall noise is clearly visible during winter in MLSY and LHC, as well as the entire monitoring period in YLTPS.

IV. DISCUSSION

A network for monitoring biodiversity represents the essential infrastructure for conservation management. By using autonomous sound recorders, a long-term soundscape monitoring network can be efficiently established. In this study, we demonstrated the application of using machine learning to analyze long-duration recordings archived in a soundscape monitoring platform. Even though the environmental noise dominates the soundscape, our analysis tools can suppress it, to investigate the structural change of biological sounds. Therefore, the remote acoustic sensing of biodiversity can be achieved by integrating machine learning and ecoacoustics.

Various animal vocalizations may be recorded in long-duration recordings. However, the lack of recognition database is the biggest challenge in acoustic-based biodiversity monitoring. In this study, we employed *k*-means clustering to investigate the complexity of animal vocalizations according to the spectral features displayed on a median-based LTS and a difference-based LTS. Although we cannot precisely identify each vocalizing species, each cluster recognizes a unique combination of vocalizing species. In this study, the measurement of bioacoustic diversity is related to the probability distribution of clusters in one day. The bioacoustic diversity reduces to 0 when only one cluster is identified. On the contrary, a higher diversity can be obtained when multiple clusters are observed in the same day. According to Fig. 5, bioacoustic diversities were higher during spring and summer due to the evident diurnal change in biological sounds, which suggested various groups of vocalizing animals, present in different diurnal periods. The clustering result not only helps us to evaluate the dynamics of biodiversity, but also reduce the effort to sort audio clips with similar acoustic features. Researchers can examine several audio clips of each cluster to

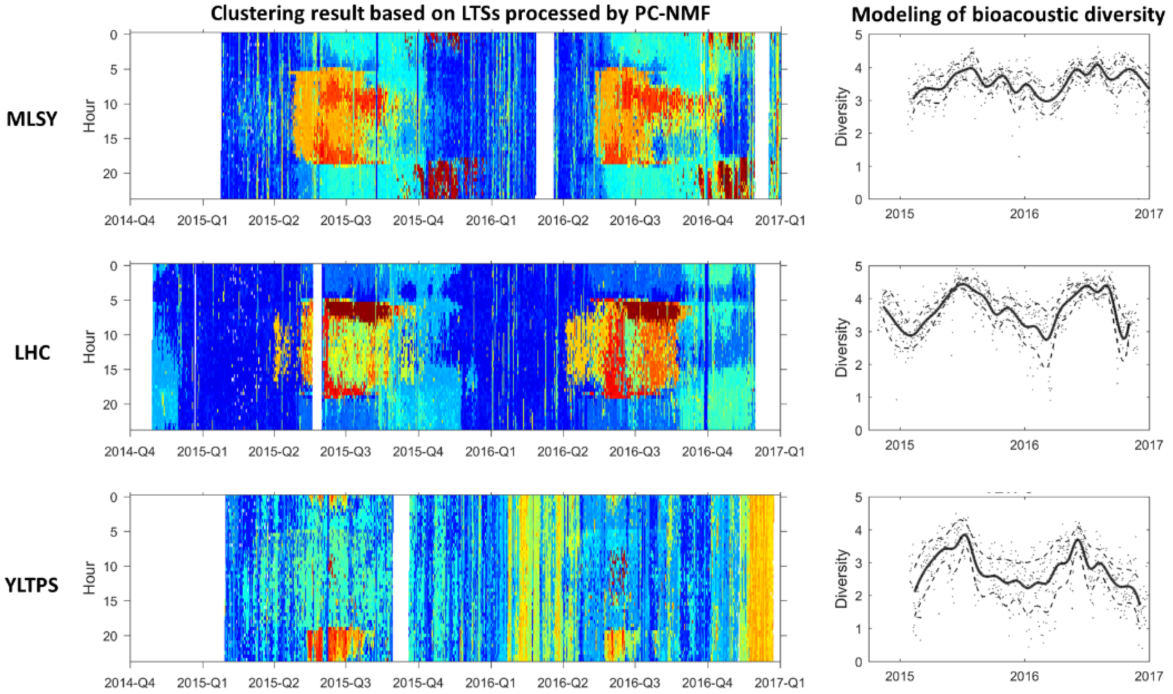


Fig. 5. Analysis results based on the biological sounds separated from median-based LTSs and difference-based LTSs. The left panels show the diurnal and seasonal change in biological sounds by using *k*-means clustering. Each color represents a different biological sound with unique spectral features. On the basis of PC-NMF, the influence of rainfall noise has been considerably reduced. The right panels show the seasonal change in bioacoustic diversity. Solid lines represent the mean diversity, dashed lines represent the standard deviation, and dots represent the diversity index measured on each day of recording.

manually identify the species composition of vocalizing animals, thus improving the acoustic-based biodiversity assessment.

Our results show that bioacoustic diversities in all three sites changed with the seasons. However, the seasonal change pattern varied based on the location. The evident seasonal change pattern at YLTPS may be because of the seasonal change of temperature in high elevation areas, with the influence of north-eastern monsoon during fall and winter. On the contrary, the weaker seasonality at MLSY may be related to a more stable climatic condition in low elevation evergreen forests. Various factors, such as altitude, latitude, weather, and many biotic and abiotic factors, can contribute to the dynamics of biodiversity. Although the detailed mechanism for the variation of bioacoustic diversity remains unclear, the long-duration recordings provide us the big data with high temporal resolution to investigate the spatio-temporal dynamics of species diversity of vocalizing animals.

In addition to the application of computing biodiversity change, the current analysis tools can also provide us information relevant to the habitat environment and weather conditions. Although we use the PC-NMF to enhance biological sounds from a noisy LTS, the difference between the separated biological sounds and original LTS is primarily associated with environmental noise (Fig. 6). From the separated environmental noise, low-frequency ambient noise was prominent at MLSY. Besides, intensive rainfall noise can be easily identified based on its broadband structure, thus making it possible to investigate the relationship between biodiversity and environmental variability, based on the various soundscape components separated from long-duration recordings.

In this study, we have shown the potential application of a long-term soundscape monitoring network in the acoustic remote sensing of biodiversity and weather conditions. We have been coordinating with local non-governmental organizations, district offices of the forestry bureau, and universities, to expand our soundscape monitoring network in Taiwan, including terrestrial and underwater environments. Currently, we are also collaborating with the Asi@Connect project, which is funded by the European Commission, to increase the capacity of acoustic-based biodiversity observation networks in ASEAN countries. The Asian Soundscape will not only serve as an open platform for archiving long-term

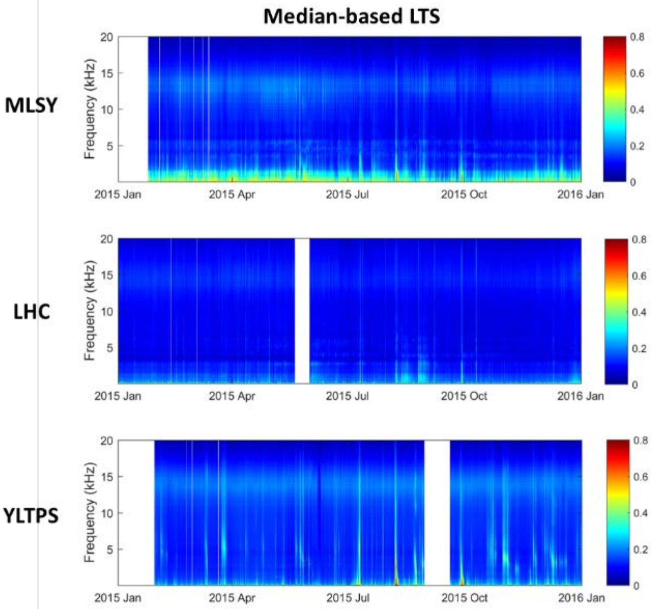


Fig. 6. Examples of environmental noise separated from median-based LTS at MLSY (upper panel), LHC (middle panel), and YLTPS (lower panel). Blank areas represent the period without recordings.

recordings and interdisciplinary collaborations, but also will provide open source tools for analyzing the acoustic data and biodiversity change. The machine learning tools for separating biological sounds from the other noise sources will facilitate the evaluation of soundscape changes. A library of wildlife sounds will also be established, based on the open soundscape database. We hope that, in the near future, through our open science platform, more researchers and citizens will be able to explore the natural soundscape and study the interactions between biodiversity, environment, and human activities.

ACKNOWLEDGMENT

This work was supported by the Ministry of Science and Technology, Republic of China (Taiwan), under a project titled "Investigation on the interactions between ecological environment, wildlife animals, and human activities using soundscape information" (MOST105-2321-B-001-069-MY3). We thank Guan-Shuo Mai (Biodiversity Research Center, Academia Sinica) for the dedicated support in maintaining the website of Asian Soundscape. We also thank all the members in Asian Soundscape for their efforts to establish recording sites and contribute soundscape recordings.

REFERENCES

[1] W. Turner, "Sensing biodiversity," *Science*, vol. 346, pp. 301–302, 2014.

[2] W. Turner, S. Spector, N. Gardiner, M. Fladeland, E. Sterling, and M. Steininger, "Remote sensing for biodiversity science and conservation," *Trends Ecol. Evol.*, vol. 18, pp. 306–314, 2003.

[3] S. Hamel et al., "Towards good practice guidance in using camera-traps in ecology: influence of sampling design on validity of ecological inferences," *Methods Ecol. Evol.*, vol. 4, pp. 105–113, 2013.

[4] C.C. Wall, P. Simard, C. Lembke, and D.A. Mann, "Large-scale passive acoustic monitoring of fish sound production on the West Florida Shelf," *Mar. Ecol. Prog. Ser.*, vol. 484, p. 173–188, 2013.

[5] C.L. Walters, A. Collen, T. Lucas, K. Mroz, C.A. Sayer, and K.E. Jones, "Challenges of using bioacoustics to globally monitor bats," in Bat Evolution, Ecology, and Conservation, R.A. Adams and S.C. Pedersen, Eds. Springer, New York, NY, 2013, pp. 479–499.

[6] T.H. Lin, H.Y. Yu, C.F. Chen, and L.S. Chou, "Passive acoustic monitoring of the temporal variability of odontocete tonal sounds from a long-term marine observatory," *PLOS ONE*, vol. 10, e0123943, 2015.

[7] J.D. Wimmer, "Acoustic sensing: roles and applications in monitoring avian biodiversity," Ph.D. dissertation, Queensland University of Technology, 2015.

[8] B.C. Pijanowski et al., "Soundscape ecology: the science of sound in the landscape," *BioScience*, vol. 61, pp. 203–216, 2011.

[9] J. Sueur and A. Farina, "Ecoacoustics: the ecological investigation and interpretation of environmental sound," *Biosemiotics*, vol. 8, pp. 493–502, 2015.

[10] K. Saito et al., "Utilizing the Cyberforest live sound system with social media to remotely conduct woodland bird censuses in Central Japan," *Ambio*, vol. 44, pp. 572–583,.

[11] M. André, M. van der Schaar, S. Zaugg, L. Houégnigan, A. Sánchez, and J. Castell, "Listening to the deep: live monitoring of ocean noise and cetacean acoustic signals," *Mar. Pollut. Bull.*, vol. 63, pp. 18–26, 2011.

[12] M.F. Baumgartner and S.E. Mussoline, "A generalized baleen whale call detection and classification system," *J. Acoust. Soc. Am.*, vol. 129, pp. 2889–2902, 2011.

[13] A. Mielke and K. Zuberbühler, "A method for automated individual, species and call type recognition in free-ranging animals," *Anim. Behav.*, vol. 86, pp. 475–482, 2013.

[14] I. Potamitis, "Automatic classification of a taxon-rich community recorded in the wild," *PLOS ONE*, vol. 9, e96936, 2014.

[15] J.F. Ruiz-Muñoz, Z. You, R.Raich, and X.Z. Fern, "Dictionary learning for bioacoustics monitoring with applications to species classification," *J Sign. Process. Syst.*, pp. 1–15, 2016.

[16] T.H. Lin and L.S. Chou, "Automatic classification of delphinids based on the representative frequencies of whistles," *J. Acoust. Soc. Am.*, vol. 138, pp. 1003–1011, 2015.

[17] J. Sueur, S. Pavoine, O. Hamerlynck, and S. Duvail, "Rapid acoustic survey for biodiversity appraisal," *PLOS One*, vol. 3, e4065, 2008.

[18] L.J. Villanueva-Rivera, B.C. Pijanowski, J. Doucette, and B. Pekin, "A primer of acoustic analysis for landscape ecologists," *Landsc. Ecol.*, vol. 26, pp. 1233–1246, 2011.

[19] N. Pieretti, A. Farina, and D. Morri, "A new methodology to infer the singing activity of an avian community: The Acoustic Complexity Index (ACI)," *Ecol. Indic.*, vol. 11, pp. 868–873, 2011.

[20] L.J. Villanueva-Rivera and B.C. Pijanowski, "Pumilio: a web-based management system for ecological recordings," *Bull. Ecol. Soc. Am.*, vol. 93, pp. 71–81, 2012.

[21] N. D. Merchant et al., "Measuring acoustic habitats," *Methods Ecol. Evol.*, vol. 6, pp. 257–265, 2015.

[22] S. Guan, T.H. Lin, L.S. Chou, J. Vignola, J. Judge, and D. Turo, "Dynamics of soundscape in a shallow water marine environment: A study of the habitat of the Indo-Pacific humpback dolphin," *J. Acoust. Soc. Am.*, vol. 137, pp. 2939–2949, 2015.

[23] T.H. Lin, S.H. Fang, and Y. Tsao, "Improving biodiversity assessment via unsupervised separation of biological sounds from long-duration recordings," *Sci. Rep.*, vol. 7, 4547, 2017.