Audio Engineering Society

Convention Paper 7893

Presented at the 127th Convention
2009 October 9–12 New York, NY, USA

Acoustics of National Parks and Historic Sites: the 8,760 hour MP3 File

Robert C. Maher

Department of Electrical & Computer Engineering
Montana State University, Bozeman, MT 59717-3780 USA
rob.maher@montana.edu

ABSTRACT

According to current U.S. National Park Service (NPS) management policies, the natural soundscape of parks and historic sites is a protected resource just like the ecosystems, landscapes, and historic artifacts for which the parks were formed. While several NPS sites have been studied extensively for noise intrusions by tour aircraft and mechanized recreation, most parks and historic sites do not yet have an acoustic baseline for management purposes. A recent initiative of the NPS Natural Sounds Office is to obtain continuous audio recordings of specific sites for one entire year. This paper explores the engineering and scientific issues associated with obtaining, archiving, and cataloging an 8,760 hour long audio recording for Grant-Kohrs Ranch National Historic Site.

1. INTRODUCTION

Visitors to U.S. National Parks expect to find unique natural features, sites of historical significance, unusual plants and animals, and open land set aside for conservation. The U.S. National Park Service Organic Act of 1916 directed the agency:

> "...to conserve the scenery and the natural and historic objects and the wildlife therein and to provide for the enjoyment of the same in such manner and by such means as will leave them unimpaired for the enjoyment of future generations." [1]

Among the features identified by the National Park Service (NPS) for protection and monitoring is the natural acoustical environment, or natural *soundscape*, of each park. The natural soundscape refers to the intrinsic acoustical environment of an area within the appropriate natural and societal context. Related terms include *natural quiet*, the *biophony* (biological sounds), the *geophony* (geological, hydrological, or meteorological sounds), and the *anthrophony* (human-caused sounds) [2, 3].

1.1. National Park Service Soundscape Policy

The NPS *Management Policies 2006* include several sections specifically addressing natural and cultural sound resources within park units [4]. The policies refer to preservation of natural sound, and, if possible, the restoration of degraded soundscapes.

- Section 4.9: Soundscape Management (*Excerpt: "The Service will restore to the natural condition wherever possible those park soundscapes that have become degraded by unnatural sounds (noise), and will protect natural soundscapes from unacceptable impacts."*) *http://www.nps.gov/policy/mp/policies.html #_Toc157232745*

- Section 5.3.1.7: Cultural Soundscape Management *(Excerpt: "The Service will preserve soundscape resources and values of the parks to the greatest extent possible to protect opportunities for appropriate transmission of cultural and historic sounds that are fundamental components of the purposes and values for which the parks were established.") http://www.nps.gov/policy/mp/policies.html #CulturalSoundscapeManagement5317*

Some NPS units, such as Grand Canyon National Park, have been the subject of extensive acoustical monitoring for regulation of the noise from air tour aircraft and other forms of mechanized recreation [5, 6]. The acoustical monitoring operations have typically involved sound surveys and short-term measurements of A-weighted sound pressure levels, following the common practice of community noise assessment.

Recently, NPS has established an initiative to obtain long-term soundscape data by making continuous audio recordings for periods ranging from several weeks to as much as an entire year. These long-term recordings will establish a baseline for scientific analysis of diurnal and seasonal variations in the natural soundscape, and for observation of long-term trends in the sonic environment of the park. It is hoped that this data will enable current and future scientists and NPS managers to use the baseline inventory for ongoing evaluation and for sustainable park management practices and guidelines.

The extreme length of these audio files—up to 8,760 hours for one year—means that it will be impossible to rely upon human listeners to audition, evaluate, and annotate the recordings. Automated processing and research in acoustical data processing and pattern recognition is essential.

1.2. Grant-Kohrs Ranch Study Site

One site currently established for long-term acoustical monitoring is Grant-Kohrs Ranch National Historic Site (GRKO), located just north of Deer Lodge, MT [7]. Grant-Kohrs is a working cattle ranch commemorating the heritage of cowboys, stock growers, and cattlemen in the history of the American West during the 19th and 20th centuries. The *cultural* soundscape associated with the working ranch is essential to visitor enjoyment and understanding. The sounds of a working ranch (bulls bellowing, draft horses pulling haying equipment, the blacksmith sharpening sickle mower blades, etc.) help immerse visitors in the historic time period the park exists to preserve.

The project underway at GRKO monitors and evaluates the natural, cultural, and community sounds that comprise the ambient acoustic environment of the historic site over the period of one calendar year. Automated acoustic instruments collect the data, including continuous digital audio recordings (MP3), calibrated Type 1 sound pressure level measurements in 1/3rd octave bands once per second, and wind speed and temperature measurements logged automatically every 10 seconds [8]. Approximately 20 GB of data are produced each month.

1.3. Outline

This paper describes the engineering and scientific issues associated with obtaining, archiving, and cataloging the 8,760 hour long audio recording for Grant-Kohrs Ranch National Historic Site, and similar considerations for future data collection sites. Among the research topics are the appropriate data collection procedures [3], the automated means to do high speed search and classification directly from the compressed MP3 data [9, 10, 11], and prospects for assessing bird, insect, and amphibian populations based on acoustic evidence [12, 13].

2. AUDIO DATA COLLECTION

Like most sites in the U.S. National Parks system, the GRKO site has never been assessed formally from an acoustical standpoint. The site has only informal, anecdotal notes, remarks, and subjective recollections characterizing the natural and cultural sounds of the park.

2.1. Purpose and Rationale

Several anticipated changes in the neighboring community of Deer Lodge may affect the visitor experience at GRKO. The Deer Lodge airport, located 2.4 km southwest of the GRKO Visitor Center, is expanding its general aviation operations, for which there is limited FAA monitoring. Highway traffic noise associated with the interstate freeway (I-90) is also increasing. The I-90 corridor runs north-south adjacent to the city of Deer Lodge, passing within 1 km of the GRKO Visitor Center. An additional impact may come from the considerable noise associated with the potential establishment of a rifle range in the vicinity of the ranch.

The GRKO staff is concerned about changes to the site's cultural soundscape due to the continued enlargement of the City of Deer Lodge airport. Recently, a large neighboring ranch was purchased, developed, and sold to individuals flying into the area in private aircraft. The airport has been expanded to accommodate this current and future land development in the area. Although some sounds from the community surrounding the park unit may not affect resources or interfere with visitor experience, substantial increases in transportation noise potentially threaten the integrity of the ranch's cultural soundscape.

In addition to the community noise threat assessments made possible by the baseline acoustical data, the long-term continuous audio recordings have the potential for extensive data mining and ecological research. For example, over 200 different bird species have been observed at the GRKO site based on historical notebooks and other records. Identifying and classifying bird calls from the long-term audio recording could help ornithologists study trends in the local bird populations and migratory patterns. Similar studies of amphibian calls, insect sounds, and other elements of the biophony can be contemplated.

2.2. Site Selection

Like most National Parks and Historic Sites, the Grant-Kohrs Ranch comprises many different soundscape zones, and therefore it would be ideal to obtain audio measurements from a large set of geographically dispersed locations. For practical reasons it was necessary to limit this particular study to a single site, based upon a variety of parameters and compromises. Because the GRKO Historic Site preserves a cultural environment rather than the natural quiet and wilderness attributes of other parks, the decision was made to choose a recording site situated adjacent to both the pasture land and the historic ranch buildings in a zone familiar to the ranch visitors who venture out beyond the visitor center. The site selection also needed to be in a location that would not be trampled by livestock or interfere with day-to-day ranch operations.

The general relationship of the monitor site and the surrounding area is shown in Figure 1.

Figure 1: Aerial view of GRKO site adjacent to Deer Lodge, MT

2.3. Equipment Deployment

The National Park Service Natural Sounds Program Office (NPS NSPO) has developed a standard procedure for long-term sound monitoring operations. The standard system is based upon a sound level meter with data logging capability, an audio recorder, and a wind speed monitor (anemometer) [8].

The sound level meter (SLM) is used to obtain the calibrated sound pressure level (SPL) in 1/3rd octave bands, recorded once per second, using an omnidirectional electret condenser measurement microphone mounted vertically approximately 1.25 meters above the ground on a tripod secured with guy wires and stakes. The analog line-level output from the

SLM is fed to the input of a digital audio recorder, such as an MP3 device, for uninterrupted recording.

The anemometer is mounted on its own tripod near the recording position. The anemometer determines the average wind speed and gust speed with a 10 second recording interval. The wind speed logging is needed to help determine if the SLM readings are ambient sound or possibly due to wind turbulence at the microphone, as described in sections 3 and 4 below.

One additional consideration is the electrical power needed to operate the SLM, audio recorder, and anemometer. The power configuration selected is a set of sealed lead-acid batteries that are recharged with a photovoltaic panel and battery charging controller. The battery capacity is sufficient to power the system for several days without sun, and the photovoltaic panel is sufficient to top off the batteries even on relatively short winter days when the sun is low in the sky.

Figure 2: National Park Service recording system deployed 2009 March 17 at Grant-Kohrs Ranch NHS (1370 meters above sea level; 46°24'27.24"N 112°44'27.60"W).

The SLM used at this site is a Larson Davis model 831, with sufficient flash memory to hold roughly 30 days of the once per second 1/3rd octave data. The sound recording system uses a Roland Edirol MP3 recorder with 32GB flash memory, providing up to 46 days of storage at a 64kbps rate. The anemometer wind speed information is logged with a HOBO brand data system with capacity sufficient for approximately 30 days. Thus, the data must be downloaded from the recording instruments at least once per month.

The SLM, recorder, and wind speed logger are placed in a sealed, weatherproof equipment case, equipped with ports for the microphone cable, anemometer cord, and power wiring. Sufficient heat is generated by the electronics to keep the case interior warm enough for operation on the cold Montana winter nights, and shade from grassy vegetation prevents overheating in the summer months.

2.4. Site Maintenance and Data Download

The monitor site requires minimal day-to-day maintenance. GRKO ranch staff visually checks the site once or twice per week to make sure the photovoltaic charging system is operating and that the tripods and guy wires have not blown over.

Approximately once per month the data logging capacity is used up and the system data must be downloaded for offline storage and analysis. The SLM data is transferred to a laptop computer using a USB memory stick, and the HOBO wind speed logging information is downloaded to the laptop via a serial cable. The month's worth of MP3 data is many gigabytes in size, so we simply dismount the 32GB flash memory chip to bring back to the lab for subsequent extraction of the data, and place an empty 32GB chip in the Edirol, ready for the next month's data.

The other monthly tasks include calibrating the SLM, replacement of desiccant in the microphone housing, examination of the cables and mounting hardware for possible damage by weather, rodents, or birds, and verifying that the batteries and charging system are in good working order. From time to time it is also necessary to swap out the SLM and microphone for laboratory recalibration and certification.

3. DATA INTERPRETATION

3.1. Aural examination

While the data collection process is still underway, we have begun to do some preliminary aural interpretation of the audio data. The typical NPS protocol is to select a 24 hour period and extract 10 seconds of sound every two minutes. This sampling process gives 720 ten second segments, or two hours of audio out of the 24 hour period. A human listener is hired to audition each of the 720 segments and note any recognizable sounds.

The primary sounds frequently identified include elements of the biophony (bird calls and insects), the geophony (primarily wind, but also thunder and raindrops), and the anthrophony (ranch vehicles, highway traffic, aircraft, nearby freight trains, ranch/domestic animals, etc.).

3.2. Visualization via spectrographic display

The 1/3rd octave sound pressure level (SPL) data provides a convenient visual way to assess the time varying spectral features of the soundscape. A trained technician can become quite adept at identifying certain classes of sound from the spectral representation, even with just 1/3rd octave data obtained once per second. Our preliminary methodology is to use the spectrographic display as a way to identify interesting segments for more careful examination and listening. In many cases this type of *directed search* is more time efficient for the aural analysis than the ten seconds per two minutes sampling approach.

An example 24-hour (midnight to midnight) spectrographic summary is shown in Figure 3. Each horizontal strip shows the frequency vs. time spectrogram for a two hour period. A windy segment is seen by the broadband stripping visible, for example, between 01:15 and 03:30. The presence of bird calls can be seen during the 06:00 through 10:20 span, noted by the presence of high frequency short-duration energy. A ranch tractor passes the monitor site approximately 13:55, and a loud train crosses the ranch site between 14:40 and 14:45, and again at 21:55.

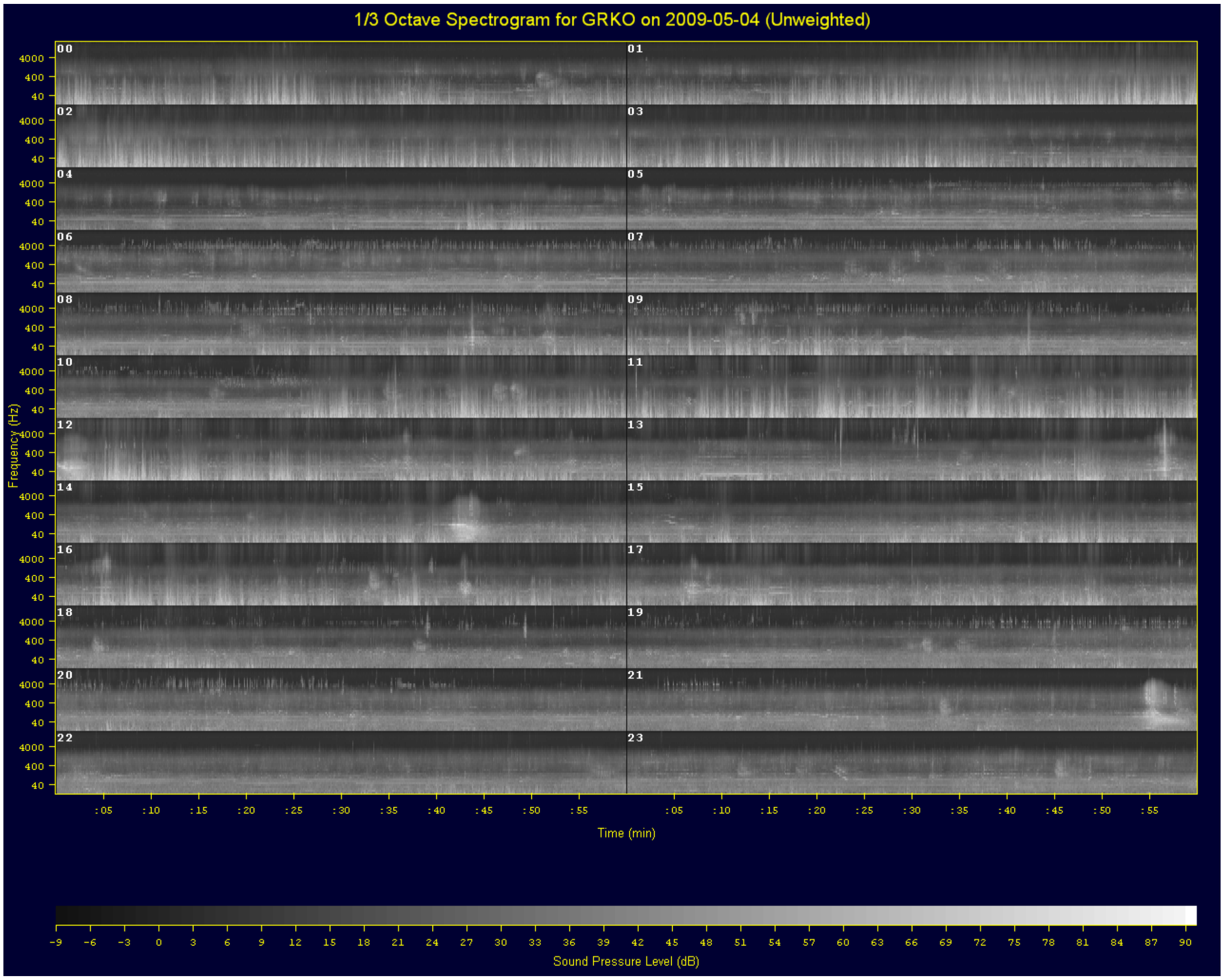

Figure 3: 24 hour spectrographic summary of the GRKO monitor site on 2009 May 4.

3.3. Effects of wind

The acoustic monitor site is located in a valley area surrounded by several mountain ranges, so wind is a nearly constant feature of the meteorological environment. Wind speeds in the 10-40 km/hr range are not uncommon, corresponding to 5 or 6 on the Beaufort wind scale.

The presence of wind is, of course, part of the natural and cultural environment of the historic site, but the air turbulence created by the wind passing over and around the microphone often leads to sufficient rumble in the microphone channel to obliterate the presence of other sounds, or even to cause microphone clipping.

The microphone is equipped with a commercial outdoor foam windscreen (32cm tall, 10cm outer diameter) and bird-repellent spike (see Figure 4), but nevertheless the wind noise can be a significant problem.

Figure 4: Windscreen (32cm x 10cm) and bird spike enclosing microphone

The standard procedure for National Park Service sound monitoring sites is to use the anemometer data as a way to determine whether the noise captured by the microphone is background sound of interest, or if it is due to wind turbulence. Wind measurements over 5 m/s indicate that the audio channel is likely to contain significant noise due to the turbulence. When NPS is performing regulatory monitoring of noise, segments with wind speeds greater than 5 m/s are eliminated from the monitor statistics.

4. AUTOMATED ANALYSIS

Clearly, practical time and cost constraints preclude that human listeners will do the complete analysis and classification of the 8,760 hours of audio obtained at even the single GRKO site, much less the anticipated data from numerous other NPS locations. Automated analysis procedures are imperative.

4.1. Feature search using SPL data

The one second 1/3rd octave SPL data provides a convenient format to detect basic features of the soundscape. Simple threshold rules can be used to uncover abrupt onsets for further examination. More elaborate classification rules are also applicable, such as identifying tonal components that can be attributable to biophonic sources, e.g., bird vocalizations.

4.1.1. Biophony

During the spring months that have been examined so far, the biological sounds that predominate are bird calls. The bird vocalizations are evident in the 1/3rd octave data as collections of narrowband components, including partials in the 1-5kHz band, with duration of several seconds or less. The spectrographic data can be used to locate candidate bird sound segments for subsequent audio processing for automated identification and classification [14, 15, 16].

In some cases, particularly the “dawn chorus” of many bird species, the density of bird calls is complicated and overlapping sounds make automated extraction difficult. Nevertheless, automated analysis of the spectrographic data should enable a trained human listener to focus on the segments with bird calls of interest.

Among the most common and repeated bird calls noted in the spring season data are the male red-winged blackbird (*Agelaius phoeniceus*), Canada goose (*Branta canadensis*), and the sandhill crane (*Grus canadensis*). See Figures 5-7 for example spectra extracted from the actual long-term audio recordings obtained at GRKO.

4.1.2. Geophony

The primary geophonic sounds during the spring months are due to wind and rain/hail. The spectrographic evidence of wind is low frequency energy with broadband extent, and this can be confirmed by comparing to the wind speed data for the corresponding time interval.

Rain detection from the 1/3rd octave record has been difficult. Rainfall has a composite sound quality in the audio recordings, including both the sound of drops hitting the microphone's windscreen and the sound of rain impacting on the adjacent ground, pasture, and ranch structures. The short time duration of individual rain drops combined with the numerous overlapping drops is not well represented by the 1-second 1/3rd octave format.

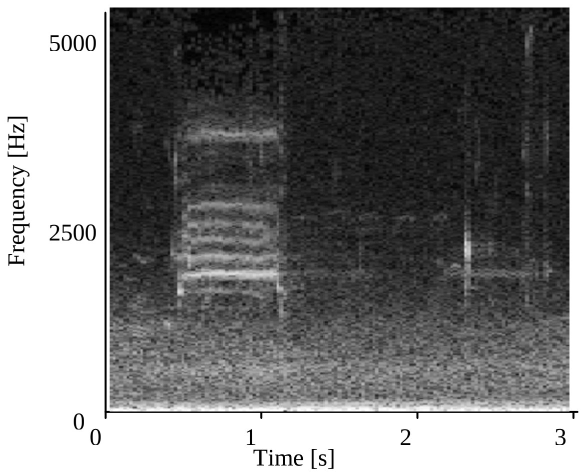

Figure 5: Bird vocalization: Red-winged blackbird (*Agelaius phoeniceus*)

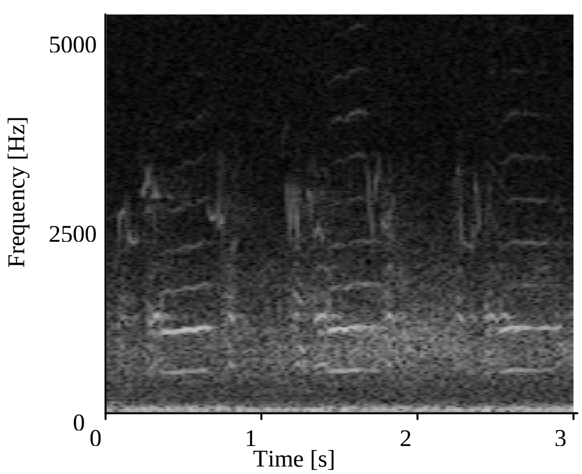

Figure 6: Canada goose (*Branta canadensis*)

4.1.3. Anthrophony

The sounds of human activity and domestic ranch animals are among the most distinctive features in the 1/3rd octave spectrogram. Some features, such as the noise from a freight train at the nearby crossing, can be identified from the relatively high sound level in the bands below 200 Hz. The train's horn is also a loud and distinctive feature.

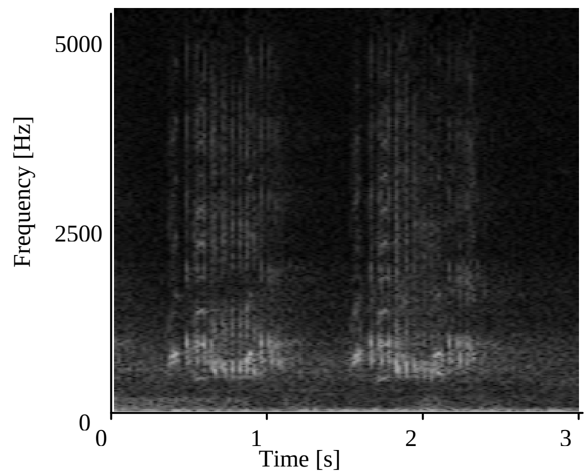

Figure 7: Sandhill crane (*Grus canadensis*)

An example one hour segment displaying the 1/3rd octave spectrogram is shown in Figure 8. The spectrographic features associated with a sequence of bird calls, a high altitude jet (turbine) overflight, and the passage of a small freight train on the tracks through the ranch are indicated in the figure.

4.2. Prospects for fast search using MP3 encoded data

In addition to the 1/3rd octave spectral information, short-time Fourier transform techniques can provide finer resolution analysis of the long-term audio recordings. Rather than resynthesizing the uncompressed audio stream and then performing the spectral analysis, the MP3 files can be examined on a frame by frame basis even without performing the full decoding operation. The subband/MDCT coefficients contain spectral information with a higher time sampling rate and finer spectral resolution than the 1 second 1/3rd octave material.

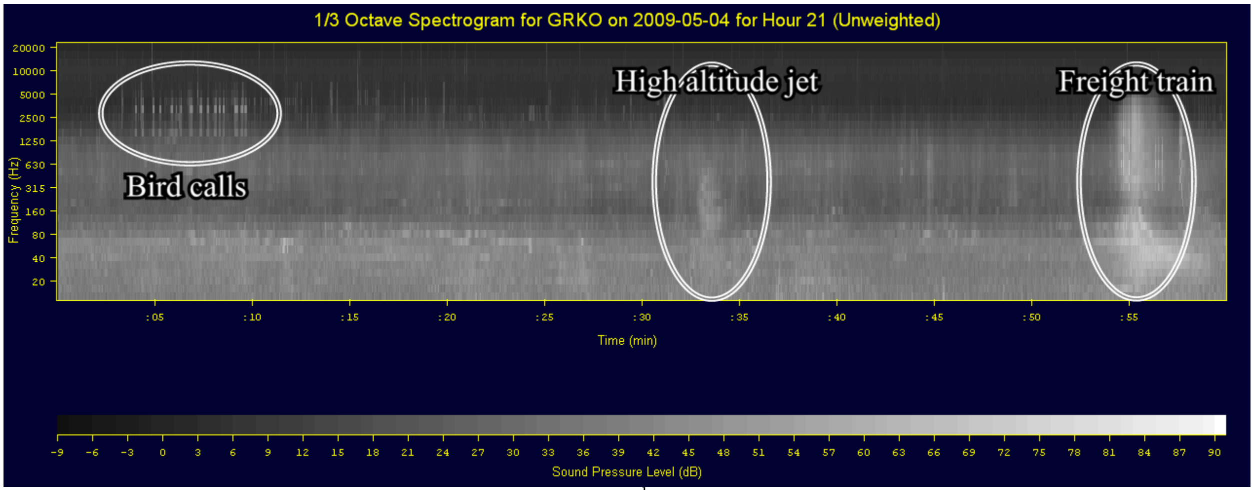

Figure 8: Example display showing the 1/3rd octave spectrogram for one hour of sound

The MP3 data blocks represent 576 audio samples, providing sufficiently good time resolution for detecting sound onsets and transitions. This allows a variety of options for pattern detection and recognition.

Prior studies have derived a set of spectral features, such as a calculation of the spectral centroid and spectral tilt, and then observed the features over time to identify changes and segmentation boundaries [8, 9, 10]. The challenge for the current project has been the need to process composite signals containing multiple simultaneous sound sources, and creating reliable and robust signal features is an ongoing area of research.

5. DOCUMENTATION AND ARCHIVING

The monthly data downloads (see section 2.4) include the 1/3rd octave SLM data (~350MB/month), the continuous MP3 data (~20GB/month), and the anemometer data (~400kB/month). In the course of a year, this single site will produce several hundred gigabytes of data that must be documented and archived.

5.1. Chronological sequencing and merging

Each of the data files includes a start time and date to facilitate sequencing. The NPS Acoustical Toolbox software can be used to organize the individual data files and to align the starting times [17]. The time sequenced and merged data is then in a format that is suitable for manual and automated processing. For example, information derived from the SLM measurements can be correlated with the wind speed data, and then the researcher can audition the actual audio recording corresponding to the interval of interest.

The documentation associated with each data collection includes detailed site information and photographs, serial numbers, equipment calibration records, and the acoustical and meteorological data files. The merged data is currently organized by monitoring site, although in the future it may be possible to interpret the data across multiple sites by calendar date or other longitudinal dimensions.

5.2. The archive problem

One of the obvious challenges of the long-term audio monitoring initiative is to house the data in a reliable, accessible manner. Even with the ongoing advances in digital storage capacity, the aggregate size of the data for one monitoring site exceeds what can casually be stored conveniently on dismountable media such as CD-ROM or DVD-ROM. The current archiving practice is to use at least two external USB disk drives with capacity of at least 500GB to hold the data and a backup copy. The repository is held offline “on the shelf” until it is needed. Online storage is certainly feasible if sufficient capacity and bandwidth can be made available.

The sustainability of the archiving procedure hinges on both the integrity of the physical storage and the commitment of the project managers to transfer data from the inevitably obsolete storage media to new storage systems on a regular basis. Of course, this problem is not unique to the long-term acoustical monitoring field, but is shared by archivists in virtually every contemporary domain.

6. CONCLUSIONS

At the time this manuscript was prepared, the GRKO project was four months into the twelve month survey period. The preliminary data are being analyzed to produce summaries of ambient sound levels in 1/3rd octave frequency bands and sparse sampling by human listeners. The summary results will characterize diurnal and seasonal variation in sound levels. The data will also be analyzed to document the audibility of identified sound sources and categories. In addition to the project documentation, and summary results, the project staff is preparing a newsletter-style document describing the project. A project overview is also available online (ece.montana.edu/rmaher/audio_monitor/grko.htm). This web site has photographs and audio data examples from the GRKO monitoring effort.

In the future, the demand for automated archiving and annotation systems will increase substantially as more and more long-term audio monitoring data is collected by the National Park Service. In the short term, the important fact is that the data are being collected and archived so that future developments in signal processing and pattern recognition can be brought to bear in mining new and useful information from the baseline data.

7. ACKNOWLEDGEMENTS

This work is supported by the Rocky Mountain Cooperative Ecosystem Study Unit (RM-CESU), Missoula, MT: Kathy Tonnessen, Natural Resources Research Coordinator, and Lisa Gerloff, Executive Coordinator.

We gratefully acknowledge the assistance and ongoing support of the Grant-Kohrs Ranch National Historic Site: Christine Ford, Integrated Resources Program Manager.

The project is conducted under the guidance and leadership of the NPS Natural Sounds Program, Ft. Collins, CO. In particular, we are pleased to acknowledge Kurt Fristrup, Emma Lynch, and Damon Joyce for sharing their time and expertise in support of this project.

8. REFERENCES

[1] National Park Service Organic Act, Public Law Chapter 408, 39 Stat. 535 et seq., 16 USC 1, 1916 Aug. 25.

[2] Krause, Bernie, "Anatomy of the soundscape: evolving perspectives," J. Audio Eng. Soc., vol. 56, no. 1/2, 2008 January/February, pp. 73-80.

[3] Maher, R.C., Gregoire, B.J. and Chen, Z., "Acoustical monitoring research for national parks and wilderness areas," Preprint 6609, Proc. 119th Audio Engineering Society Convention, New York, NY, October, 2005.

[4] U.S. National Park Service, *Management Policies 2006*, U.S. Government Printing Office, 2006 (http://www.nps.gov/policy/MP2006.pdf)

[5] U.S. National Park Service, "Report on Effects of Aircraft Overflights on the National Park System," United States Department of the Interior Report to Congress pursuant to P.L. 100-91 (The National Parks Overflights Act of 1987), Report NPS-D-1062 (NTIS Number: PB95-263323), July 1995.

[6] Miller, N.P., "US National Parks and management of park soundscapes: A review," Applied Acoustics, vol. 69, 2008, pp. 77-92.

[7] Maher, R.C. "Baseline sound monitoring plan for Grant-Kohrs Ranch national historic site," J. Acoust. Soc. Am., vol. 125, no. 4, part 2, (abstract), 2009, p. 2716

[8] U.S. National Park Service, "Standard Operating Procedure for Acoustic Monitoring," Natural Sounds Program, Fort Collins, CO, 2008.

[9] Yapp, L., and Zick, G., "Speech recognition on MPEG/Audio encoded files," Proceedings IEEE International Conference on Multimedia Computing and Systems '97, Ottawa, Ontario, Canada, June 1997, pp. 624-625.

[10] Tzanetakis, G. and Cook, P., "Sound analysis using MPEG compressed audio," IEEE International

Conference on Acoustics, Speech and Signal Processing, Istanbul, vol. 2, 2000, pp. II761-II764.

[11] Pfeiffer, S. and Vincent, T., "Survey of compressed domain audio features and their expressiveness," Proc. SPIE, vol. 5021, 2003, pp. 133-147.

[12] Rand, A.S., and Drewry, G.E., "Acoustic monitoring at fixed sites," in W.R. Heyer, M.A. Donnelly, R.W. McDiarmid, L.-A.C. Hayek, and M.S. Foster, eds., *Measuring and monitoring biological diversity: standard methods for amphibians*. Smithsonian Institution Press, Washington, D.C., 1994, pp. 150-152.

[13] Barichivich, W.J., "Guidelines for building and operating remote field recorders (automated frog call data loggers)," Appendix IV. in C. K. Dodd, Jr., *Monitoring Amphibians in Great Smoky Mountains National Park*, U.S. Geological Survey Circular 1258, 2003.

[14] Chen, Z., and Maher, R.C., "Semi-automatic classification of bird vocalizations using spectral peak tracks," J. Acoust. Soc. Am., vol. 120, no. 5, 2006, pp. 2974-2984.

[15] Heller, J., and Pinezich, J., "Automatic recognition of harmonic bird sounds using a frequency track extraction algorithm," J. Acoust. Soc. Am., vol. 124, no. 3, 2008, pp. 1830-1837.

[16] Brandes, T.S., "Feature Vector Selection and Use With Hidden Markov Models to Identify Frequency-Modulated Bioacoustic Signals Amidst Noise," IEEE Trans. Audio, Speech, and Language Processing, vol. 16, no. 6, 2008, pp. 1173-1180.

[17] Joyce, D., "Acoustic Monitoring Toolbox User Guide," U.S. National Park Service, Natural Sounds Program, Fort Collins, CO, 2009.