Baseline Sound Monitoring at Grant Kohrs Ranch NHS

Final Report
7 September 2012

Grant-Kohrs Ranch National Historic Site (GRKO), located just north of Deer Lodge, Montana, 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.

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 physical 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 and Night Skies Office is to obtain continuous audio recordings of specific sites for one entire year. Accordingly, GRKO staff arranged to conduct a baseline acoustical survey at Grant-Kohrs Ranch to monitor and evaluate the natural, cultural, and community sounds that comprise the ambient acoustic environment of the historic site over the period of one calendar year. This report summarizes the engineering and scientific findings associated with the acquisition, archiving, and cataloging of the 8,760 hour baseline audio survey for Grant-Kohrs Ranch National Historic Site, conducted 17 March 2009 to 19 March 2010.


Background

The Grant-Kohrs Ranch was established to provide an understanding of the frontier cattle era. Congress believed the best format for visitor understanding was to maintain the site as a working ranch, with all the sights, sounds, and sensations associated with ranching. Thus, 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 National Park Service *Management Policies 2006* include several sections specifically addressing natural and cultural sound resources within park units.

- 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*

GRKO currently has no data characterizing the natural and cultural sounds of the park. Several anticipated changes in the neighboring community of Deer Lodge may affect the visitor experience at GRKO. The Deer Lodge airport, located 1.5 miles 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 I-90 freeway is also increasing. The I-90 corridor runs north-south adjacent to the city of Deer Lodge, passing within 0.7 of a mile 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.

An imminent threat to the site's cultural soundscape is 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 jets and helicopters. 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.

Scope of Work

The approved research plan for this project comprised the following elements:

1. Obtain digital recordings and detailed sound pressure level readings to document sound sources and associated sound levels in primary acoustic zones in the park.

2. Analyze these data to provide concise summaries that will assist park managers in evaluating the effects of park activities and extrinsic sound sources on the acoustical environment. This will help park staff follow guidelines set forth in 2006 management policies for soundscape management. (*See 2006 Management Policies Sec. 4.9-Soundscape Management, 5.3.1.7-Cultural, 8.4-Overflights and Aviation Uses, and 8.2-Visitor use*).

3. Produce data which can be used to inform park management plans.

4. Disseminate the procedures and technical findings in the appropriate audio engineering professional literature as a basis for "best practices" standardization.

5. Enable new research regarding semi- or fully-automated analysis and unsupervised segmentation of soundscape recordings.

6. Increase the skilled workforce for soundscape acoustics by educating student technicians in the proper techniques for acoustical data processing and analysis.

The remainder of this report provides a summary description of the accomplishments toward each of the task elements. An appendix containing example reports and papers is also included.

Task 1: Data Acquisition Plan and Implementation

GRKO staff worked with Montana State University (MSU) and the National Park Service Natural Sounds and Night Skies Division (NSNSD) to perform a long-term (12 month) assessment of the characteristic seasonal and diurnal soundscape at a site located approximately 200 meters west-southwest of the historic Ranch House complex, situated between the Visitor Center and the Clark Fork River (see Figures 1 and 2).

Figure 1: GRKO ranch orientation detail.

Automated acoustic instruments provided by the NSNSD were deployed at the site to collect the bulk of the data (see Figure 3). The measurements included calibrated 1 second L_{eq} 1/3rd octave sound levels obtained with a Larson Davis 831 sound level meter (ANSI Type 1 certification). A calibrated digital audio (MP3) recording was collected continuously using the analog audio output feed from the sound level meter and a Roland Edirol audio recorder operating at 44.1kHz sample rate, encoding the data at 16-bit, 64kbps data rate. Wind speed measurements were logged automatically every 10 seconds using a HOBO anemometer and digital data logging system.

The sound level meter and audio recorder were powered from an external 12V battery system powering 5 volt and 3 volt DC regulators. The batteries were kept charged using a photovoltaic panel and an electronic battery charge control system. The estimated current for the meter and the audio recorder was under 1 ampere on average, so four 12-amp-hour rechargeable batteries were used to provide several days of power even if insufficient sunlight was available for a day or two due to cloudy skies on short winter days.

The HOBO logger was powered by 4 AA alkaline cells, which would provide at least 8 weeks of powered operation.

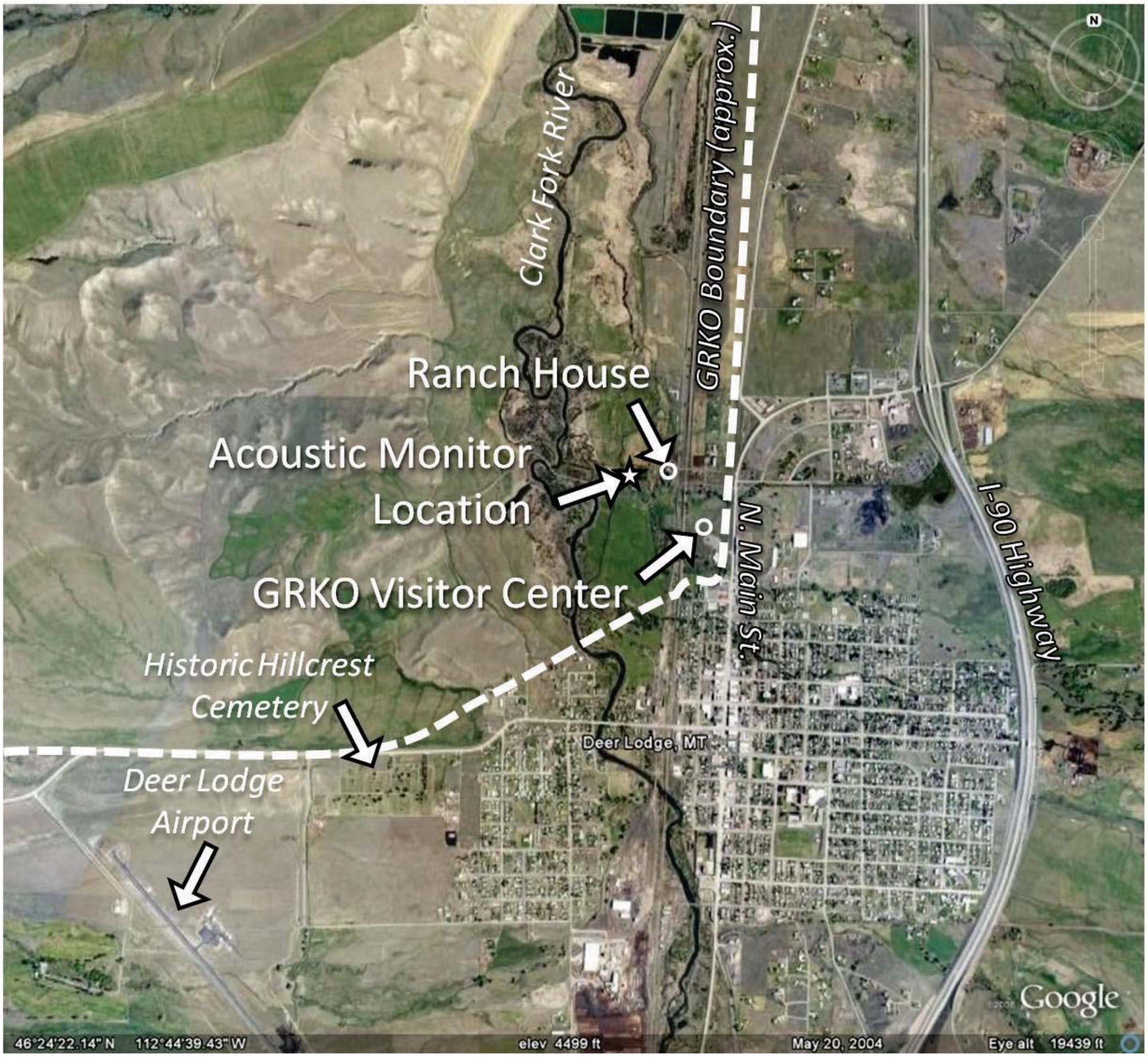


Figure 2: GRKO ranch and neighboring community.

Figure 3: Acoustical monitoring site equipment deployed at GRKO on 17 March 2009.

With the photovoltaic system providing continuous power, the data collection period between servicing visits to the site was limited by the storage capacity of the instruments.

- The SLM storage capacity was 2GB. Each 1 second L_{eq} 1/3rd octave sound level measurement required approximately 148 bytes, plus the file storage overhead, indicating that the 2GB capacity would provide just over 5 months of storage.
- The HOBO logger 10 second wind speed storage provided more than 30 days of data capacity. As noted above, the battery life for the logger was estimated to be at least 8 weeks.
- The Edirol MP3 recorder was equipped with a 32GB flash storage chip. At the data rate of 64kbps, the capacity was approximately 46 days.

Thus, the limitation on site service was set principally by the wind speed logger at just over 30 days.

The initial deployment of the system occurred on Tuesday, March 17, 2009. The site visits occurred on the following dates:

Visit Date	Days Since Prior Visit
3/17/2009	--
4/10/2009	24
5/5/2009	25
5/28/2009	23
6/22/2009	25
7/17/2009	25
8/8/2009	22
9/5/2009	28
10/3/2009	28
10/24/2009	21
11/15/2009	22
12/12/2009	27
1/7/2010	26
1/30/2010	23
2/21/2010	22
3/19/2010	26

At each site visit, the following steps were performed.

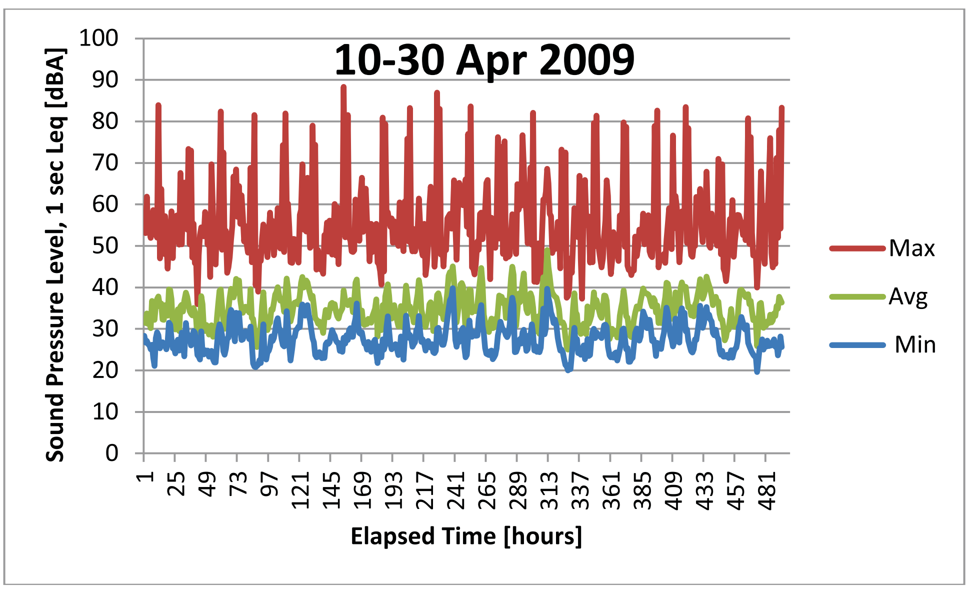
- Examine site and equipment; look for any visible evidence of changes or damage.
- Give an audible time slate, indicating the date and time of the visit.
- Examine the photovoltaic system, charge controller, and measure the battery voltage.
- Open the equipment case and verify that the instruments have power and are running properly.
- Fill out the Site Visit Data Sheet.

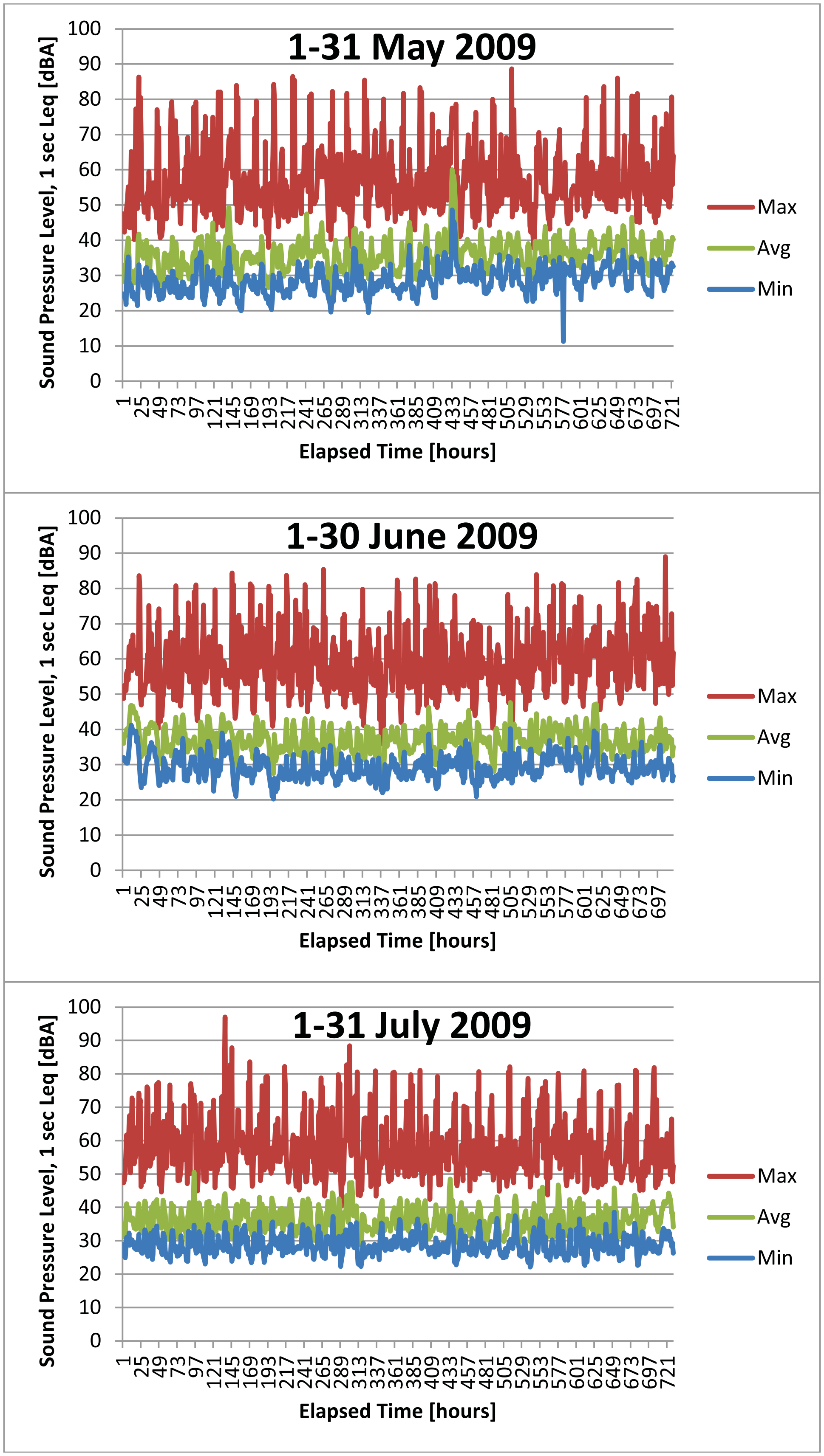
- Stop the SLM and the Edirol. Save the SLM data onto a flash memory stick, and swap the 32GB flash chip in the Edirol for a new, empty chip.
- Attach the HOBO logger to a laptop computer and download the wind speed data.
- Using the laptop, reset the HOBO and launch for the next data collection.
- Clear the SLM, reset its clock, and reset the unit for the next data collection.
- Set the Edirol clock, verify its settings, and reset for the next data collection.
- Check the dessicant level in the microphone housing, and replace if necessary.
- Attach the calibrator to the microphone and recalibrate the SLM. Also record the calibration tone with the Edirol to get a calibration level in the audio data stream.
- Remove the calibrator and reattach the microphone wind screen.
- Verify that the recording instruments are running, then close and lock the equipment case.
- Give an audible time slate, indicating the date and time of the visit.
- Secure the equipment and leave the site.

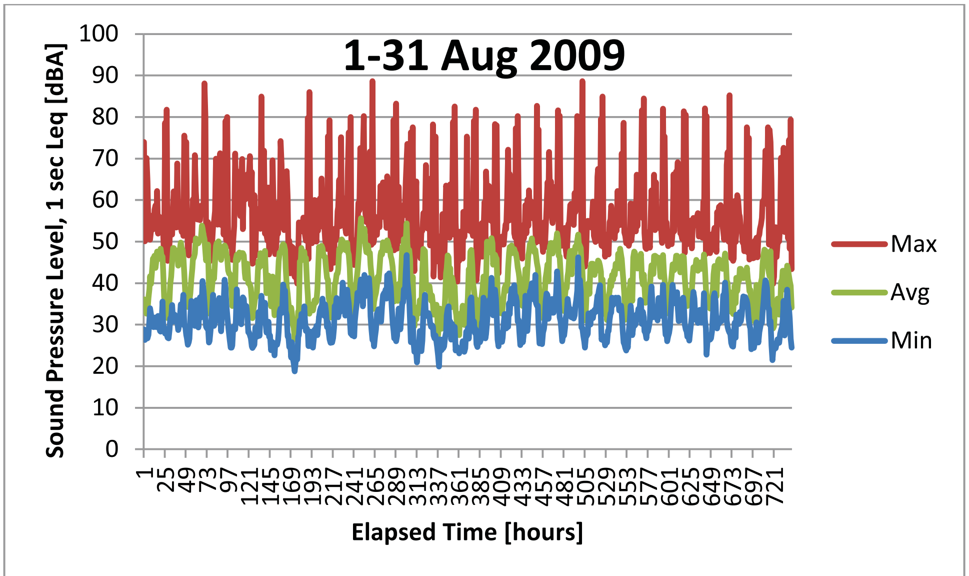
On two occasions, April 10, 2009 and January 30, 2010, the SLM was found to be malfunctioning. The SLM apparently had a software glitch of some kind that caused the system to lock up and stop recording the 1/3 octave data. Other than these two issues, the system appeared to function as designed, even during periods of extreme cold, heat, ice, and wind.

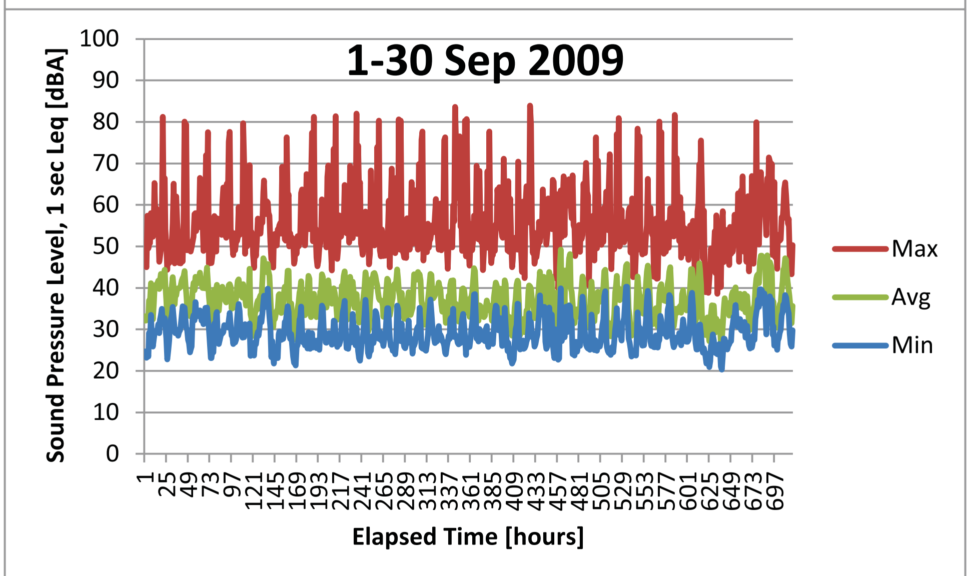
Task 2: Summary Sound Level Information

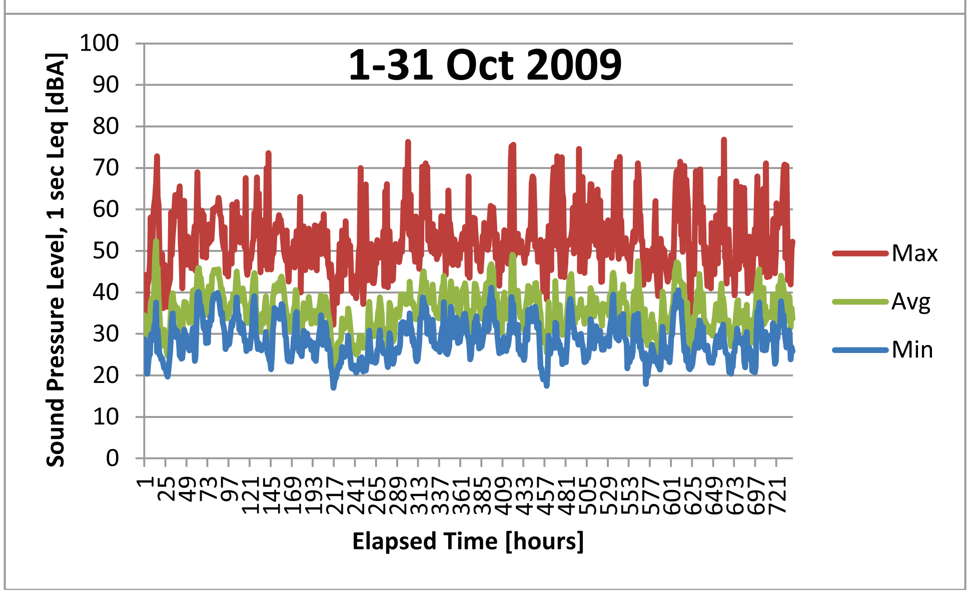
The SLM data for overall sound level measured once per second were processed to determine the maximum, minimum, and average level for each hour. The information excludes the periods for which the SLM did not record data properly (17 Mar - 9 Apr 2009 and 13 - 30 Jan 2010).

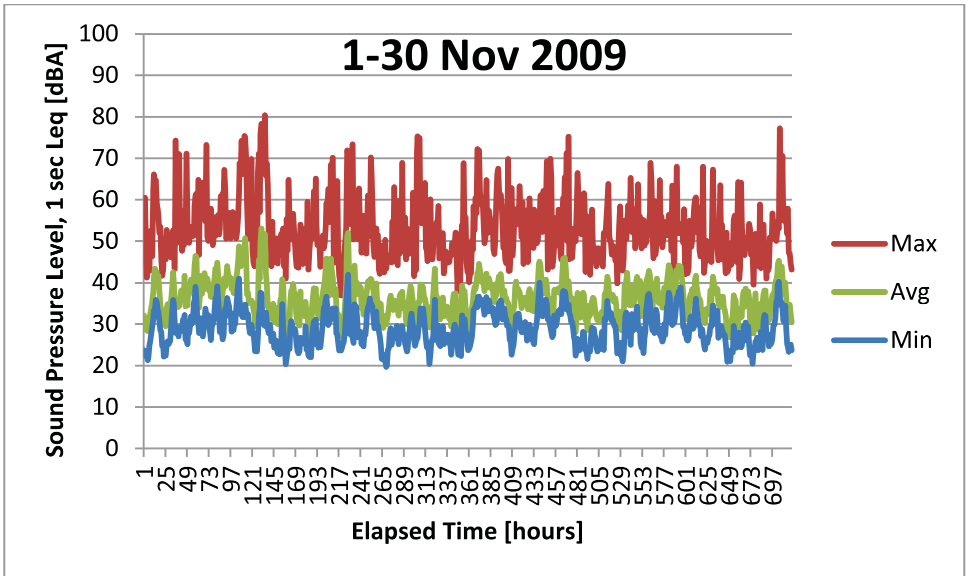


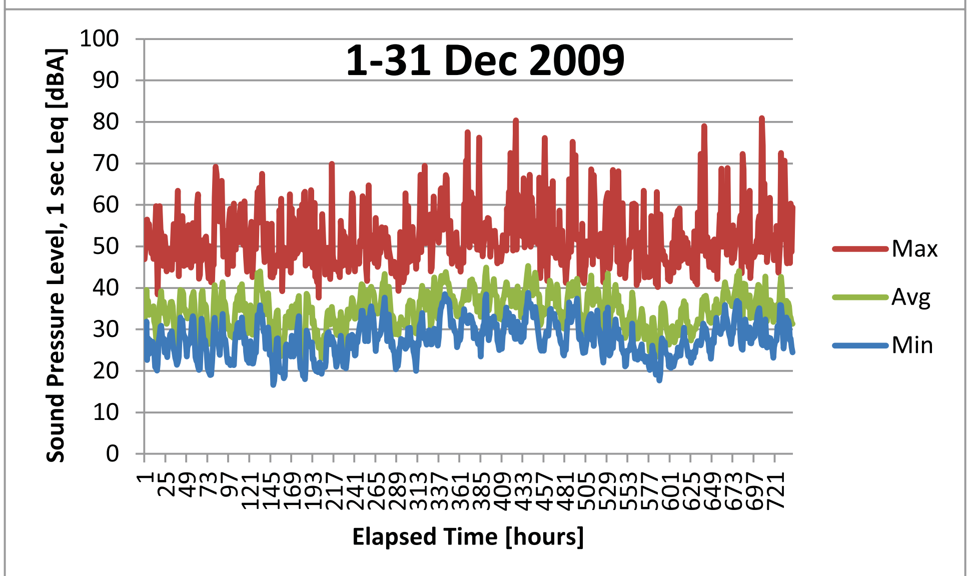

1-31 May 2009
Sound Pressure Level, 1 sec Leq [dBA]
Elapsed Time [hours]
Max
Avg
Min
1-30 June 2009
Sound Pressure Level, 1 sec Leq [dBA]
Elapsed Time [hours]
Max
Avg
Min
1-31 July 2009
Sound Pressure Level, 1 sec Leq [dBA]
Elapsed Time [hours]
Max
Avg
Min

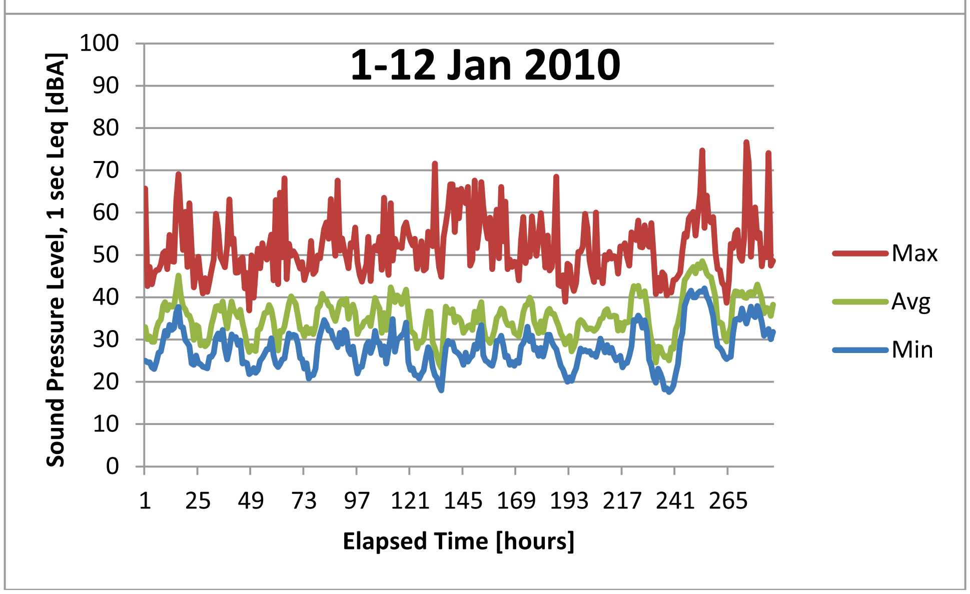

1-31 Aug 2009
Sound Pressure Level, 1 sec Leq [dBA]
Elapsed Time [hours]
Max
Avg
Min

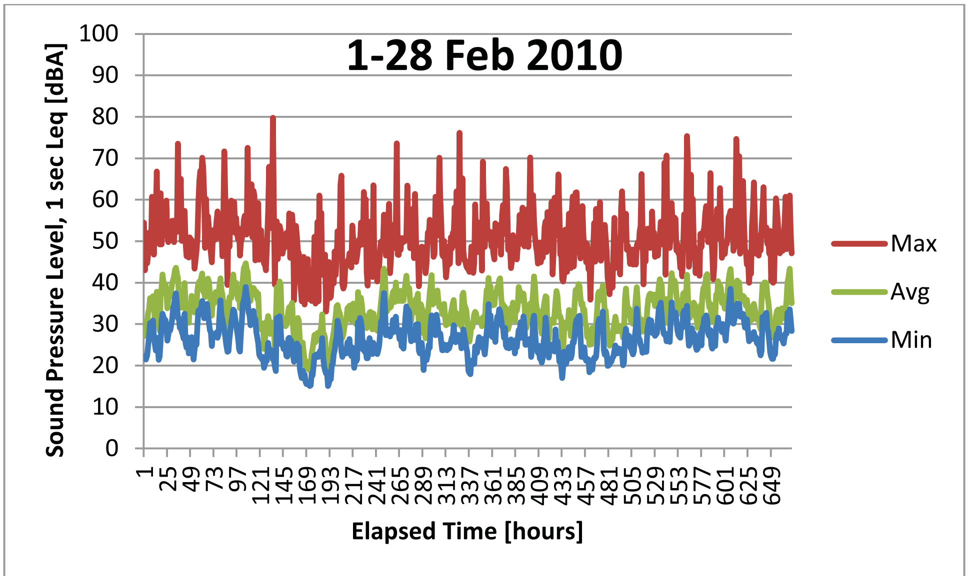

1-30 Sep 2009
Sound Pressure Level, 1 sec Leq [dBA]
Elapsed Time [hours]
Max
Avg
Min

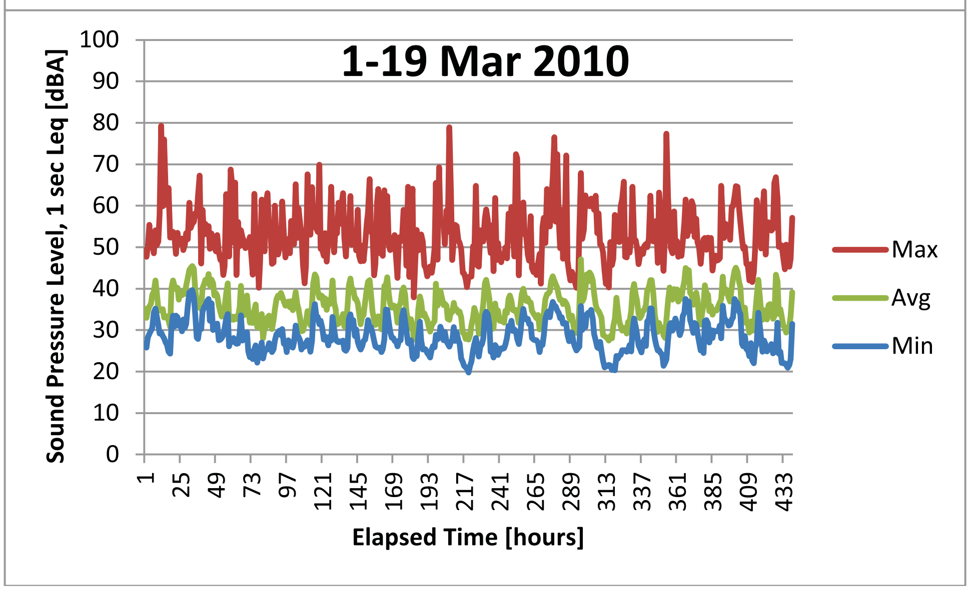

1-31 Oct 2009
Sound Pressure Level, 1 sec Leq [dBA]
Elapsed Time [hours]
Max
Avg
Min


1-30 Nov 2009
Sound Pressure Level, 1 sec Leq [dBA]
Elapsed Time [hours]
Max
Avg
Min


1-31 Dec 2009
Sound Pressure Level, 1 sec Leq [dBA]
Elapsed Time [hours]
Max
Avg
Min


1-12 Jan 2010
Sound Pressure Level, 1 sec Leq [dBA]
Elapsed Time [hours]
Max
Avg
Min





The minimum, maximum, and average sound pressure level information, as well as the MP3 audio recordings, provide several important objective and subjective observations.

- The data frequently show a diurnal pattern in which the sound level is higher during the daytime and then lower during the evening and overnight. The increase in daytime level is attributable to human activity (traffic sounds, railroad, ranch machinery, livestock), birds, and the effects of wind, which tend to be greater during the day.
- The data show a seasonal variation, with greater average sound levels during the summer months and lower sound levels during the winter. The average level during the winter months is often below 40 dBA, while the average level during the summer is frequently above 40 dBA.
- The maximum sound level during most months is under 80 dBA, but levels exceeding 90 dBA are common in certain summer months. The highest overall level (over 95 dBA) was recorded in July, and was found to be due to a very loud thunder clap.
- Selective listening to the MP3 audio data reveals the general acoustical characteristics of the GRKO site:
    - Audible highway noise from nearby I-90 is present virtually all of the time.
    - During the recording year, considerable railroad traffic occurred on the tracks passing north-south through the ranch itself. The rail traffic was primarily local trains ferrying contaminated

soil from the Milltown Dam site near Missoula to the storage site south of Deer Lodge in the Anaconda area.
  - The traditional cultural soundscape of the ranch is audible, although usually in the subtle form of cattle sounds, trucks and tractors, roosters crowing, horse whinnies, and similar sounds. In some sections of the recording sounds from the community of Deer Lodge are heard, including a siren (fire signal?), alarms, truck back-up beeps, and similar sounds.
  - Bird sounds, particularly in the dawn and dusk phases of the spring, summer, and fall, are very distinctive throughout the recorded data.
  - Wind noise is present in many segments of the recording. In some cases the wind causes microphone "clipping" or dropouts.

Task 3: Data Relevance to Park Management

As an important cultural location among the slate of U.S. National Historic Sites, Grant-Kohrs Ranch provides visitors with a unique soundscape comprising wildlife, livestock, human voices and activity, ranch vehicles, wind and rain, and other sounds inherent to GRKO's charter. The Historic Site managers are directed by NPS Management Policy to understand and to preserve the cultural soundscape. In the NPS 2006 Management Policy document, section 5.3.1.7 dealing with cultural soundscapes, states that "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."

The data collected from the year-long audio survey provides a baseline for the managers' use. Specifically, the managers may use the data to assess the impact of future development and proposed land use changes in the Deer Lodge area.

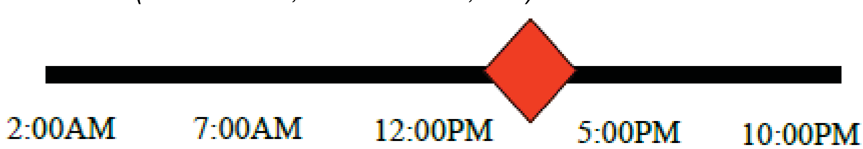
For example, at some point in the future a proposed change to general aviation services at the Deer Lodge airport could be assessed for the additional aircraft overflights and aircraft noise for GRKO visitors, compared to the levels found during the survey year. Similarly, potential changes to sound levels at the Historic Site due to activity at the nearby Powell County fairgrounds could be estimated and evaluated.

Another important potential benefit of the acoustical survey data is in the evaluation of wildlife at the GRKO site, particularly the diversity of birds and other sound-producers, such as coyotes, insects, and possibly frogs and other animals. Park managers may use the audio recordings to assess the range of dates on which particular bird species are heard, and compare that information to future survey data.

The student technicians who assisted with this project (see Task 6 below) have suggested several ways in which they would recommend GRKO visitors to learn about the soundscape and to utilize the audio data. Here are a few of their suggestions.

Sound Data Slider

Visual display shows a bell-curve graph of sound in certain intervals. (Intervals can be several hours, weeks, months, etc.). Visitor can move the slider to different intervals and see how the sound changes throughout that interval. (Hour to hour, week to week, etc.).

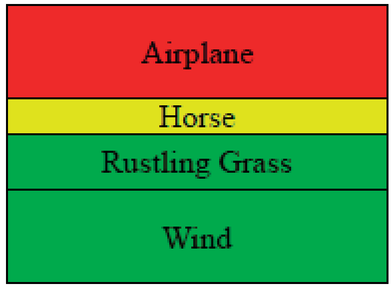


Sound Identification:

Strip out other sounds so the visitor only hears one sound. (Horse, bird call, goose, etc.).Then, slowly add in the other background sounds (train, cars on a road, wind, any other background sound) until the visitor can no longer hear the original sound. Show how other sounds can mask the original sound.

What is the sound at Grant-Kohrs made up of?

Let the visitor see what the sound at Grant-Kohrs is made up of. Visually have lights or a chart that make up a vertical bar. Show in the bar the composition of the sounds in the park. This can also be done by percentage breakdown in the bar depending on the level of the sound. (Airplane is loud, horse is quiet.)



How does Grant-Kohrs compare with other places?

Show a day (or any interval of time) at Grant-Kohrs visually and compare it to some other well known places. (Grand Canyon, Yellowstone, New York City). Let the visitor see and also listen to the difference.

Task 4: Dissemination of Procedures and Findings

Among the important effects of the GRKO acoustical survey has been the dissemination of results to professionals in the fields of acoustics and audio engineering.

The first dissemination task was the creation of a public website describing the project goals, status, and example findings. The site includes several links to MP3 audio recordings demonstrating the variety of sounds present at the ranch during the monitoring period. The URL for the website is currently: http://ece.montana.edu/rmaher/audio_monitor/grko.htm

The website text and images is included in the Appendix of this report.

The second and ongoing dissemination task has been the presentation of project information at meetings of the appropriate professional organizations, including the Acoustical Society of America (meetings in Portland, Baltimore, Seattle, and San Diego), and the Audio Engineering Society (meetings in New York, San Francisco, and Denver).

- R.C. Maher, "Baseline sound monitoring plan for Grant-Kohrs Ranch national historic site," invited paper, J. Acoust. Soc. Am., vol. 125, no. 4, part 2, p. 2716 (abstract), Portland, OR, 2009.
- R.C. Maher, "Acoustics of national parks and historic sites: the 8,760 hour MP3 file," Preprint 7893, Proc. 127th Audio Engineering Society Convention, New York, NY, October 2009.
- R.C. Maher, "Cultural soundscape of the Grant-Kohrs Ranch national historic site," invited paper and soundscape concert, J. Acoust. Soc. Am., vol. 127, no. 3, part 2, p. 1745 (abstract), Baltimore, MD, 2010.
- R.C. Maher, "Maintaining sonic texture with time scale compression by a factor of 100 or more," Preprint 8250, Proc. 129th Audio Engineering Society Convention, San Francisco, CA, November 2010.
- R.C. Maher, "Automated analysis and interpretation of long-term soundscape audio recordings," invited paper, J. Acoust. Soc. Am., vol. 129, no. 4, part 2, p. 2570 (abstract), Seattle, WA, April, 2011.
- R.C. Maher and Christine Ford, "Soundscape collaboration for science, management, and public outreach at a national historic site," invited paper, J. Acoust. Soc. Am., vol. 130, no. 4, part 2, p. 2497 (abstract), San Diego, CA, October 2011.

- R.C. Maher and J. Studniarz, "Automatic search and classification of sound sources in long-term surveillance recordings," Proc. Audio Engineering Society 46th Conference, Audio Forensics—Recording, Recovery, Analysis, and Interpretation, Denver, CO, June 2012.

Copies of the presentations that required a published paper are included in the Appendix.

The third dissemination effort has been informal and public presentations for special interest groups meeting at Grant-Kohrs Ranch itself, and at other community venues. It is hoped that these types of events will continue in the future. Montana State University is committed to a variety of outreach efforts, and continued cooperative work with GRKO is planned.

- R.C. Maher, "Baseline Sound Monitoring at Grant-Kohrs Ranch National Historic Site," invited workshop presentation, National Cooperative Ecosystem Study Unit (CESU) Coordinators Meeting, Deer Lodge, MT, May 2009.
- R.C. Maher, “Sound monitoring of parks and nature areas, including Grant-Kohrs Ranch,” presentation for GRKO Youth Conservation Core (YCC) students, Deer Lodge, MT, June 2009.
- R.C. Maher, "The Science of Sound: Acoustics and Audio Engineering," invited seminar presentation, Aspen Pointe Speakers Series, Bozeman, MT, September 2010.
- R.C. Maher, “Soundscape management: The impact of change on the soundscape,” presentation for Grant-Kohrs Teacher Workshop, Deer Lodge, MT, June 2012.
- R.C. Maher, “The sounds of silence and the sounds of rivers,” invited presentation, Montana Institute on Ecosystems seminar series, Missoula, MT and Bozeman, MT, September 2012.

Task 5: Enable Semi- or Fully-Automated Analysis of Soundscape Recordings

A long-held goal of the audio signal processing community is referred to as *computational auditory scene analysis*, or CASA. A CASA system would be able to process an audio recording and identify the various sound sources, sound characteristics, and other aspects of the environment in which the recording was made. At present, the capabilities of automated analysis systems remains rudimentary--especially when compared to the abilities of trained human listeners when performing audio analysis tasks.

The challenge for human listeners is to maintain vigilance and attention. This is especially true if one contemplates having human listeners study the more than 8,000 hours of audio material produced in the GRKO study. It would literally require listeners the equivalent of 365 days, 24 hours per day, to listen to all of the recordings. The cost of compensating the listeners for their time would also be prohibitive.

Although in the scope and budget of this relatively small project it was not conceived to create a fully automated analysis system, there were several important steps in this direction. First, the ability to create the summary sound pressure level (SPL) charts shown earlier in this report is very feasible with the sound level meter data. Second, the use of spectrographic display techniques, including the features of the National Park Service audio monitor toolkit software, is fully compatible with the data collected in this project. Third, we have worked on several approaches to do high-speed search of the MP3 data looking for particular sound types, and these preliminary results have been described in our recent paper for the AES Audio Forensics Conference in Denver this past summer (2012). Finally, we have studied several approaches that allow the potential for rapid audition of long audio recordings using what we describe as *time-lapse aural display* (see the AES paper from 2010).

As of the writing of this report, there are ongoing audio signal processing research projects at Montana State University that use the GRKO sound monitor data as a vital and unique part of the investigation. The additional and ongoing work will form the basis of future publications and innovations in this field.

Task 6: Educating Student Technicians

Because Montana State University is an educational institution, and because of the ongoing need for research and development in soundscape principles, it was considered very important that this project serve to ncrease the skilled workforce for soundscape acoustics by educating student technicians in the proper techniques for acoustical data processing and analysis.

Three students were engaged during the data collection, processing, and ongoing analysis phases of this project.

Darrin Reed was an electrical engineering student working on his Masters Degree during the preliminary portion of this project. Darrin received his MS degree in 2009, served as an intern during spring 2010, and since that time has continued his education in audio signal processing at the University of Oldenburg in Germany. Darrin was involved in this project by developing the data extraction and display techniques we have used in many of our experiments. Darrin also helped with the site visits during the summer of 2009. Darrin has excellent prospects for continued work in audio signal processing throughout his career.

Shruti Sah was an electrical engineering student working on her Bachelor of Science Degree during the data collection portion of this project. Shruti, an international student from India, received her BS degree in December, 2010. While a student, she was employed as an hourly research assistant to help process and organize the audio, SLM, and windspeed data. Shruti was particularly interested in the environmental aspects of the soundscape studies, and she will likely seek out ways to apply her engineering expertise in the environmental technology field.

Joe Studniarz is a current computer engineering students working on his BS Degree. He is expected to complete his studies in the spring of 2013. Joe has been extremely effective in his software development work as an hourly research assistant. Joe performed the data processing necessary to resolve several minor issues of missing or misinterpreted SLM data, and has also worked extensively in techniques to use the compressed MP3 data to try out some of our research ideas about finding information directly in the compressed file, rather than having to reconstruct the audio signals for subsequent processing. Since the funding from the GRKO project was used up last year, Joe has been supported by miscellaneous funding from internal sources at Montana State University.

Appendix

Published papers, presentations, and website information.

AUTOMATIC SEARCH AND CLASSIFICATION OF SOUND SOURCES IN LONG-TERM SURVEILLANCE RECORDINGS

ROBERT C. MAHER AND JOSEPH STUDNIARZ

Department of Electrical and Computer Engineering, Montana State University, Bozeman, MT USA
rob.maher@montana.edu

The increasing availability of forensic audio surveillance recordings covering days or weeks of time makes human audition impractical. This paper describes the rationale and potential application of several techniques for high-speed automated search and classification of sound sources and sound events in long-term forensic audio recordings. Methods that can operate directly on perceptually compressed bitstreams without full decoding are of particular interest. Example applications include identification of aircraft overflights, the presence of human speech utterances, gunshot sounds, and other forensically relevant audio signals.

INTRODUCTION

An emerging problem in audio forensics involves the virtual explosion of data made available by long-term surveillance and soundscape audio recordings [1-5]. Long-term audio surveillance recordings may contain speech information and also non-speech sounds such as environmental noise, audible warning and alert signals, footsteps, mechanical sounds, gunshots, and other acoustic information of potential forensic interest. Manually searching and auditioning hours and hours of audio material to detect subtle, telltale sound events can be both tedious and error prone.

In this investigation we assess several techniques for high-speed automated search and classification of sound sources and sound events in long-term forensic audio recordings. The fundamental objective of these techniques is to transform the audio signal into a domain in which the desired signal attributes can be identified and classified.

The most promising current techniques use some form of short-time spectral analysis to localize in time, amplitude, and frequency the signal parameters of interest, and then to match these parameters to the classification dimensions. Although many techniques have been applied to this problem, the need for improved identification and classification techniques remains an open research question.

1 SEGMENTATION AND INTERPRETATION

Prior work that is relevant to forensic analysis of long-term surveillance recordings comes from several different research areas. These include techniques to process an audio recording to detect the presence of speech or music, distinguish between speech and non-speech sounds, recognize a particular speaker or song, and identify portions of the recording that have attributes of interest for subsequent manual investigation or follow-up [6-14]. Insights and techniques are also borrowed from the field of computational auditory scene analysis [15].

Three example systems and applications are representative of the prior research in this field. In 2005, Härmä, et al. [1], reported on an experiment with an automatic acoustic surveillance system that used a microphone to monitor the acoustical environment in an office for a two month period. The system identified “interesting” acoustical events and recorded them for additional processing. The basic detection and segmentation process used an adaptive spectral model to track the slowly-varying background noise profile, and an acoustic activity detection algorithm based on the departure of the currently measured spectral snapshot from the background noise profile. The acoustic events indicated by the detection algorithm were then analyzed to create a feature vector consisting of parameters such as RMS value, spectrum centroid, duration, and bandwidth. A k-means clustering algorithm was applied to the event parameters to classify acoustic events based on their parameter similarity.

A system described in 2010 by Wichern, et al. [10], used a variety of derived parameters, such as loudness, spectral centroid, and harmonicity, to enable classification and potentially identification of environmental sounds in continuous audio recordings. Like Härmä, et al., the Wichern experiment relied upon a parameterized representation to enable clustering and classification, but unlike Härmä’s background spectral profile, Wichern, et al., calculated the parameters continuously for the entire audio stream, and used the observed changes in those parameters to infer the onset and end of the acoustic events. Wichern, et al., also experimented with a framework to retrieve audio events based upon user query information.

Also in 2010, we developed a procedure to identify changes in *sonic texture* as an indication of audio events for an application involving extreme time-scale compression of long-term audio recordings [16]. Our definition of sonic texture was the time-variant fluctuation in the 1/3rd octave band levels. The method used a one second time interval and a 1/3rd octave spectral average to capture the time-variant spectral character of the signal. The event criterion was to monitor a threshold change in one or more of the 1/3rd octave bands. For each 1 second frame, the process examined the next several frames to determine any repetitive fluctuations. The result was a map of the textural transitions, with a goal to retain the time segments with a high number of sonic texture transitions, at the expense of the segments with lower activity.

2 SPECTRAL TEMPLATE CONCEPT

The insights derived from these existing techniques have led us to consider a time-frequency orientation in which two separate two-dimensional filters, or templates, are constructed. One template is designed so that it preferentially selects spectral components that are narrow in frequency but relatively broad in time, corresponding to tonal or quasi-harmonic content, while the other 2-D filter is designed to pass spectral components that are broad in frequency but relatively narrow in time, corresponding to impulsive, abrupt, and other similarly brief events. This approach is intended to uncover coherent acoustic information in the presence of incoherent broadband noise, and is in several ways an extension of our prior work on forensic audio quality enhancement [17]. The current spectral implementation uses similar data handling and formatting procedures to the audio enhancement system described before.

2.1 Application principles

The key features of the proposed audio identification and classification system are:

- Implementation of temporal and spectral filters that operate in the time-frequency domain.
- Description of the sonic events in terms of time-frequency templates, treating the spectrographic information as an image processing task.
- Use of multiple time-frequency resolutions in parallel to match the processing resolution to the time-variant signal characteristics.

These features are intended to mimic the approach used by human forensic examiners when considering long-term surveillance data. The examiner can benefit by examining a spectrographic (frequency content vs. time) display to look for features and patterns worthy of detailed examination.

2.2 Example framework

A spectrographic representation of one hour of audio recorded outdoors in a semi-rural area is shown in Figure 1. The recording contains a plethora of sound sources that overlap in time and in frequency. There are frequent bird vocalizations and the sounds of domestic animals, mechanical sounds from vehicles, and a variety of sounds attributable to wind.

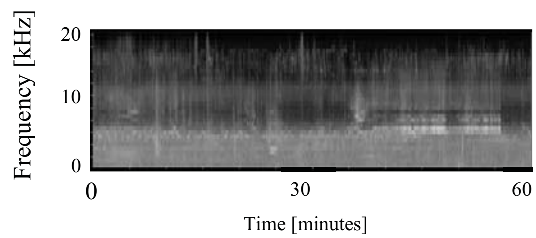


Figure 1: Spectrographic representation.

The spectrogram was created using the 1/3rd octave filterbank with 1 second average spectra. The resolution of each “pixel” in the spectrogram is therefore 1 second in width by 1/3rd octave in height. The white and light gray portions indicate time-frequency intervals in which the spectral energy is high, while the dark gray and black regions indicate low energy. The horizontal textural bands indicate steady-state or quasi-stationary sound sources, while the light colored vertical lines indicate impulsive sounds (relatively short duration with relatively broad frequency extent).

In Figure 1, consider the event visible as a bright interval at approximately 36 minutes in the 4-10kHz range. That sound source corresponding to that distinctive feature is probably not immediately recognizable from the spectrographic information alone, but additional context information and analysis can be employed to suit a particular investigation. In this case, the particular sound is an overflight by a single-engine piston-powered aircraft. The subsequent bright areas between 40 and 50 minutes are a power lawnmower’s internal combustion engine starting up, followed by the sound of the mower moving with the spinning blade engaged.

Figure 2 shows an enlarged time scale spectrographic representation of the aircraft overflight and the beginning of the lawnmower event from ~36 minutes into the record shown in Figure 1.

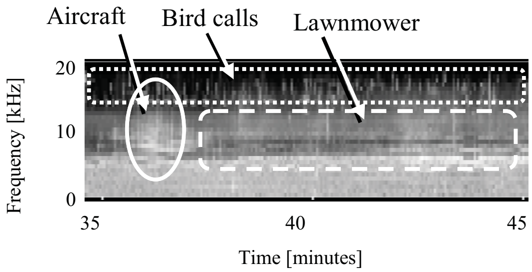


Figure 2: Enlarged portion of the spectral segment shown in Figure 1.

Decomposing the spectrographic representation of Figure 1 and Figure 2 can be thought of as an image processing exercise. For example, we can be interested in locating the vertically-oriented *edges* in the spectrographic image. The edges can be interpreted as the onsets and ends of sonic events as they evolve with time. Applying a high pass filter to the each of the 1/3rd octave filter output sequences produces the magnitude display shown in Figure 3, where the steepest detected edges are indicated as light gray and white.

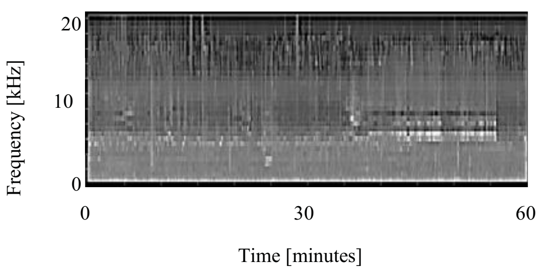


Figure 3: Spectral change "edge" representation.

The temporal edges can be further refined to enhance particular onset requirements. We currently determine the most appropriate detection settings for a particular recording by manual iteration. Fully automating the event/edge detection selections is a key part of our ongoing research.

2.3 Application considerations

Two fundamental issues arise when applying this spectral method. First, the time scale aspects must be adjusted to provide a meaningful level of detection. For example, the onset information depicted in Figure 3 includes speckle texture due to the quasi-stationary noise present in the recording, which can tend to obscure narrowband events of short duration in noisy recordings.

The second fundamental issue is the way in which sound events with a very gradual amplitude envelope appear from the spectrographic viewpoint. The proposed spectral method is best suited for sonic events with a sharp onset, and events that occur gradually, such as a high altitude aircraft approaching from a distance, do not provide a suitable "edge" in the 2-D framework. Developing appropriate techniques to reveal information regarding low-level sound events and sounds with slow amplitude envelope onset and/or release remains a research topic.

2.4 Implementation Optimization

The work on this identification and classification procedure has been conducted via software written in the C language and in Matlab. We have deliberately avoided any hardware-specific implementations up to this point because the research is driven by algorithmic flexibility and experimentation rather than time and efficiency constraints. A fully functional system would clearly require fast algorithms and implementation details that would allow processing at many times faster than real time. There is great potential for using graphics-oriented signal processing hardware in this regard.

A related development has been to use perceptually compressed audio bitstreams as the front-end for the analysis process without the need to decode the audio waveform itself [18]. Perceptual compression systems such as MP3 produce a bitstream containing a time-variant spectral analysis of the original signal and scale factors for the encoded audio, and this information can be extracted from the bitstream and interpreted without synthesizing the audio material itself. Since the 2-D filters used in the proposed identification and classification system operate in the frequency vs. time domain, allowing the initial processing to occur on the spectral parameters extracted from the perceptual coder bitstream could lead to greater potential efficiency compared to a process involving reconstructing the compressed bitstream into a time signal, only to have to perform a subsequent spectral analysis again on the decoded audio.

3 CONCLUSIONS

The availability of long-term audio surveillance recordings presents both opportunities and challenges for the field of audio forensics. Research to extract meaningful and forensically useful information from long-term recordings via automated processing remains an essential goal. The work described in this paper includes several features that will be a useful basis for future comprehensive solutions.

Many audio forensic investigations operate with an official query, such as "Are there any gunshots present in the recording between time *X* and time *Y*?" In these cases it would be desirable to have a suite of gunshot-related templates to use for the identification and classification task, and the proposed method will potentially be very useful. Similarly, in cases where the request is to determine if there are any distinctive sounds in a certain long segment of a recording, an automated method to identify candidate sound events can save time and potentially improve the examiner's efficiency and reliability. Nevertheless, in other cases the audio forensic examiner may be asked a more general question, such as "Can you help identify any of the audible background sounds between time *X* and time *Y*?" In response to such a query, the use of automated methods may be of less applicability due to the highly subjective nature of the query.

REFERENCES

[1] A. Härmä, M.F. McKinney, and J. Skowronek, "Automatic surveillance of the acoustic activity in our living environment," *IEEE Int. Conference on Multimedia and Expo*, 4 pp. (2005).

[2] D.P.W. Ellis and K. Lee, "Accessing Minimal-Impact Personal Audio Archives," *IEEE Multimedia*, vol. 13, no. 4, pp. 30-38 (2006).

[3] J.P. Ogle and D.P.W. Ellis, "Fingerprinting to Identify Repeated Sound Events in Long-Duration Personal Audio Recordings," *Proc. ICASSP*, Honolulu, HI, pp. I-233—I-236 (2007).

[4] Krause, B., "Anatomy of the soundscape: evolving perspectives," *J. Audio Eng. Soc.*, vol. 56, no. 1/2, pp. 73-80 (2008).

[5] R.C. Maher, "Acoustics of national parks and historic sites: the 8,760 hour MP3 file," *Proc. 127th Audio Engineering Society Convention*, New York, NY, Preprint 7893, (2009).

[6] E. Wold, T. Blum, D. Keislar, and J. Wheaten, "Content-based classification, search, and retrieval of audio," *IEEE Multimedia*, vol. 3, no. 3, pp. 27-36 (1996).

[7] J. Foote, "Content-based retrieval of music and audio," Multimedia Storage and Archiving Systems II, *Proc. of SPIE*, vol. 3229, pp.138-147 (1997).

[8] L. Zhu, Y. Wang, and T. Chen, "Audio feature extraction and analysis for scene segmentation and classification," *Journal of VLSI Signal Processing Systems for Signal, Image, and Video Technology*, vol. 20, no. 1/2, pp. 61-79 (1998).

[9] B.J. Gregoire and R.C. Maher, "Map seeking circuits: a novel method of detecting auditory events using iterative template mapping," *Proc. IEEE Signal Processing Society 12th DSP Workshop*, Jackson Lake, WY, pp. 511-515, (2006).

[10] G. Wichern, J. Xue, H. Thornburg, B. Mechtley, and A. Spanias, "Segmentation, Indexing, and Retrieval for Environmental and Natural Sounds," *IEEE Trans. on Audio, Speech, and Language Processing*, vol. 18, no. 3, pp. 688-707 (2010).

[11] S. Davies and D. Bland, "Interestingness Detection in Sports Audio Broadcasts," *Ninth Int. Conference on Machine Learning and Applications*, Washington, DC, pp. 643-648 (2010).

[12] G. Pietila, G. Cerrato, and R.E. Smith, "Detection and identification of acoustic signatures," *Proc. 2011 NDIA Ground Vehicle Systems Engineering and Technology Symposium*, Dearborn, MI, (2011).

[13] A. Muscariello, G. Gravier and F. Bimbot, "An efficient method for the unsupervised discovery of signalling motifs in large audio streams," *9th Int. Workshop on Content-Based Multimedia Indexing (CBMI)*, Madrid, Spain, pp. 145-150 (2011).

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

[15] A. Bregman, *Auditory Scene Analysis: The Perceptual Organization of Sound*, MIT Press, 1990.

[16] R.C. Maher, "Maintaining sonic texture with time scale compression by a factor of 100 or more," Preprint 8250, *Proc. 129th Audio Engineering Society Convention*, San Francisco, CA (2010).

[17] R.C. Maher, "Audio enhancement using nonlinear time-frequency filtering," *Proc. Audio Engineering Society 26th Conference, Audio Forensics in the Digital Age*, Denver, CO (2005).

[18] G. Tzanetakis and P. Cook, "Sound analysis using MPEG compressed audio," *Proc. ICASSP*, Istanbul, Turkey, pp.II761-II764 (2000).

Automatic Search and Classification of Sound Sources in Long-Term Surveillance Recordings

Robert C. Maher and Joseph Studniarz

Electrical and Computer Engineering

Montana State University - Bozeman

MONTANA STATE UNIVERSITY | College of ENGINEERING

Mountains & Minds

Outline

- Introduction
- Experimental Procedure
  - Example recordings
  - Feature detection
- Experimental Results
  - Time-variant spectra
  - Identifying events
  - Effects of temporal and spectral overlap
- Discussion and Conclusions

Introduction

- Long-term audio recordings
  - Continuous recordings of urban areas and natural soundscapes
  - Duration may be days, weeks, months, …
  - Human audition infeasible due to length
- Forensic applications
  - Detection and Classification
  - Event reconstruction
  - Timeline Assessment

Introduction (cont.)

- Source may be deliberate surveillance or inadvertent surveillance
  - Wildlife studies
  - Regulatory monitoring
  - Acoustic surveys
  - Electronic newsgathering
  - Amateur A/V recordings

Introduction (cont.)

- Prior research in speech/non-speech segmentation, music recognition, and computational auditory scene analysis
- General research questions:
  - Search for a target sound, or classify all sounds?
  - Will conventional pattern matching techniques do the job?
  - What level of performance is necessary?

Experimental Procedure

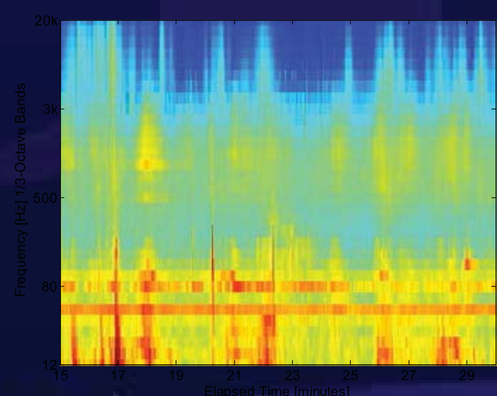
- Our example recordings
  - Grant-Kohrs Ranch (2009-10): 8,700 hours
  - Nyack River experiment (2012): 29 hours
- Sound composition
  - Biophony—birds, frogs, insects, mammals
  - Geophony—wind, rain, hail, thunder, waterfall
  - Anthrophony—aircraft, automobiles, trains, domestic sounds

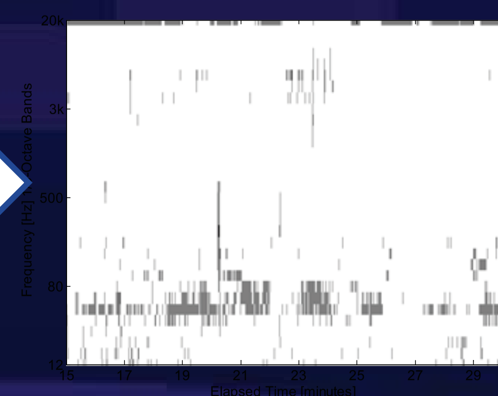
Experimental Proc. (cont.)

- Typical recordings contain a background texture punctuated by foreground sounds—but also gradual evolution and overlaps in time and frequency.
- Isolated pattern matching not applicable in general.
- Current compromise: use automated search to identify sections of interest for subsequent audition

Technical Approach

- Treat spectrographic information as an *image*, or as a 3-D *surface*
- Sound events comprise *edges*, *ridges*, *cliffs*, and *valleys* in the spectrogram
- Plan: process the spectrographic information to reveal the feature space





Spectral-domain Filters

- Vertical edges = broad in frequency, narrow in time: clicks
- Horizontal edges = narrow in frequency, broad in time: tones
- Diagonal edges = tonal sweeps
- “Waffle” pattern viewpoint

Forensic Example

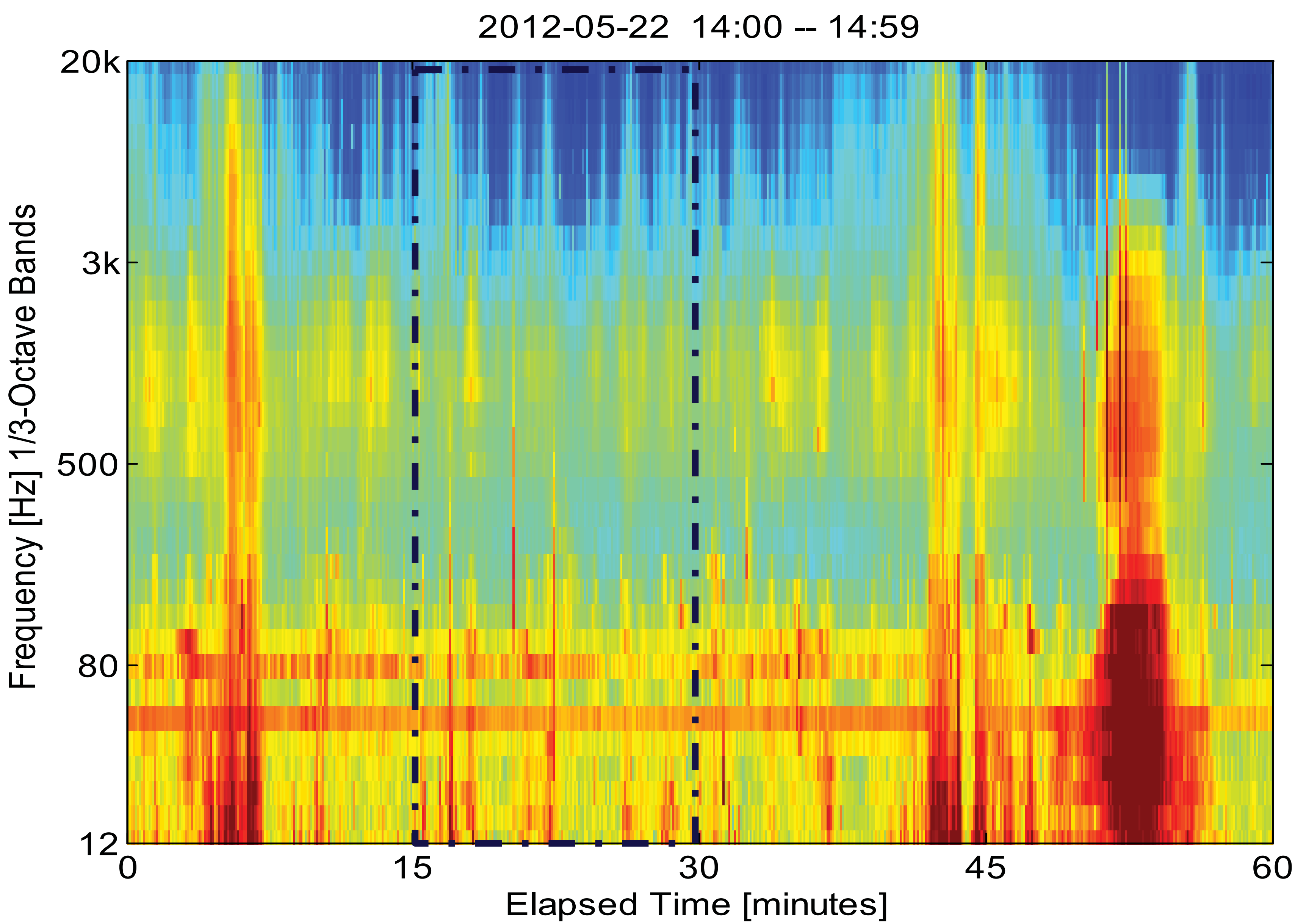
- One hour audio recording
  - Question 1: Is the sound of a gunshot present at any time in the recording?

  If the answer to question 1 is *yes*,
  - Question 2: Is there more than one recorded shot?

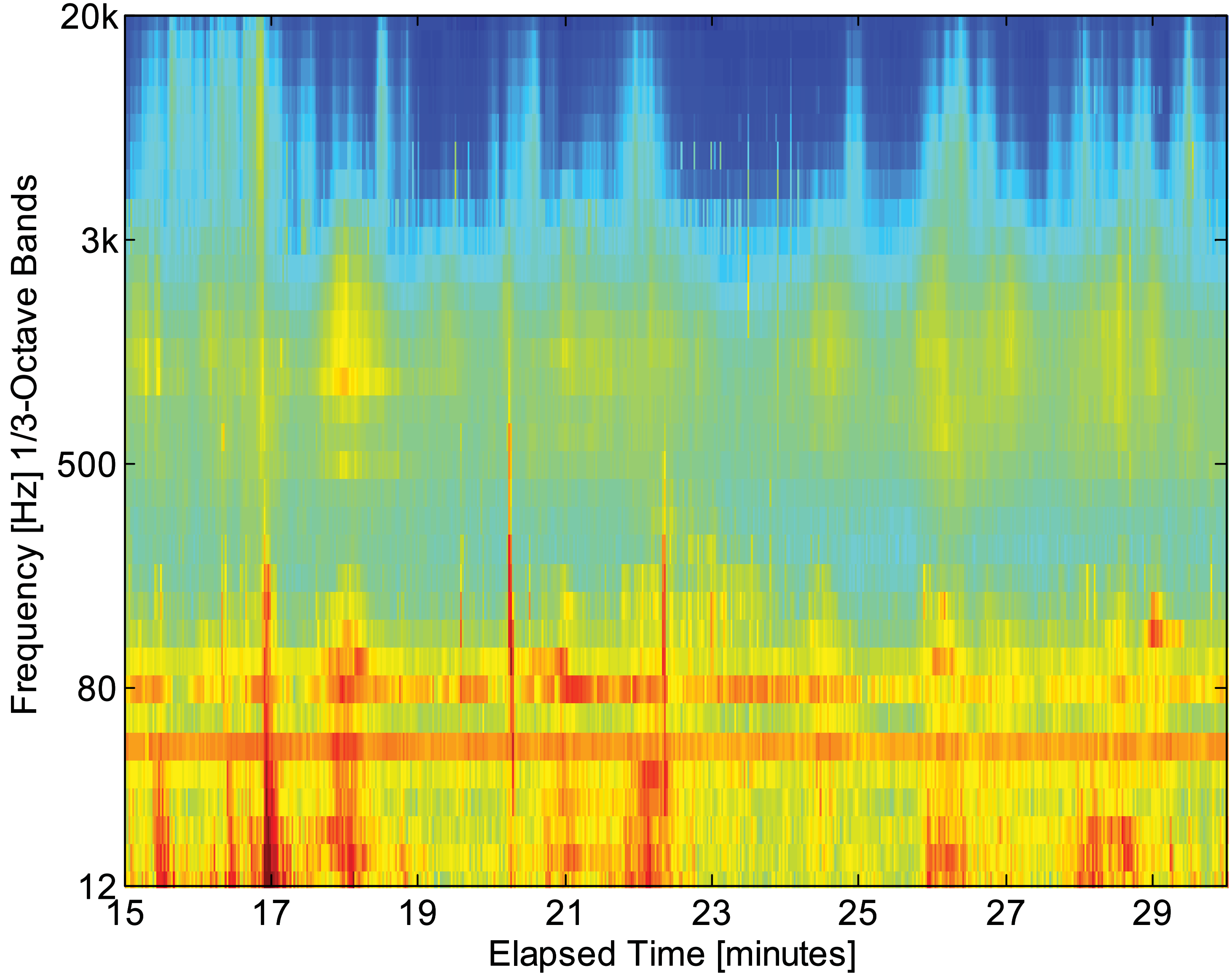
  If the answer to question 2 is *yes*,
  - Question 3: Are the multiple shots from the same or from different firearms?

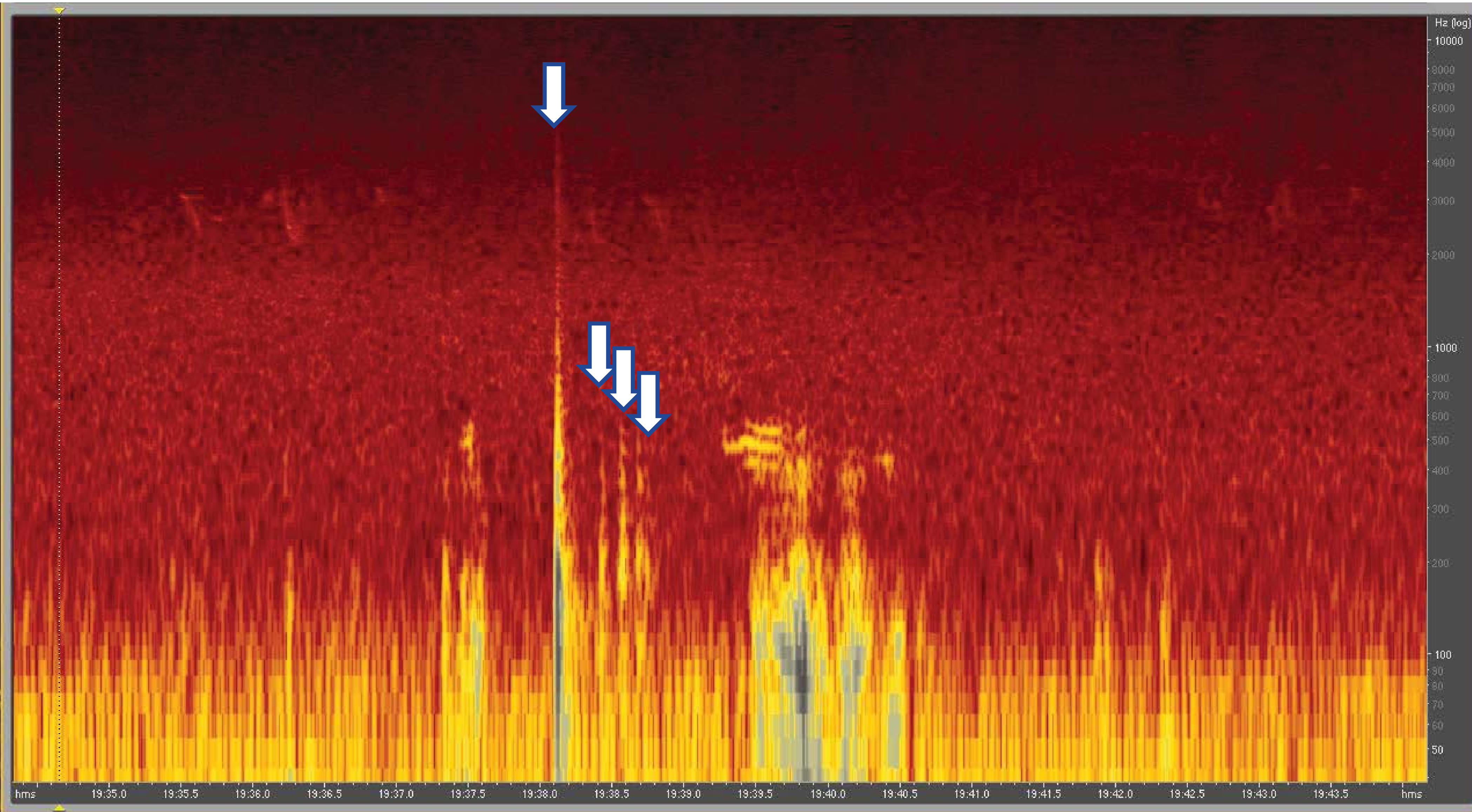
Forensic Example (an aside)

- This is a “needle in the haystack” analogy
- Finding the needle may be extremely difficult
- If no needle is found, it may mean:
  - no needle was present, or
  - the search process missed it
- But if a needle IS found, it will generally not be misinterpreted as some other object



2012-05-22 14:00 -- 14:59
Frequency [Hz] 1/3-Octave Bands
20k
3k
500
80
12
0
15
30
45
60
Elapsed Time [minutes]


20k
3k
500
80
12
Frequency [Hz] 1/3-Octave Bands
15
17
19
21
23
25
27
29
Elapsed Time [minutes]


Hz (log)
10000
8000
7000
6000
5000
4000
3000
2000
1000
800
700
600
500
400
300
200
100
90
80
70
60
50
hms
19:35.0
19:35.5
19:36.0
19:36.5
19:37.0
19:37.5
19:38.0
19:38.5
19:39.0
19:39.5
19:40.0
19:40.5
19:41.0
19:41.5
19:42.0
19:42.5
19:43.0
19:43.5
hms

Experimental Results (cont.)

- Edge-based spectro-temporal search can be implemented easily
- Good for abrupt sound events or narrow-band tones
- Sounds with spectral and temporal overlap are not handled well

Future Work

- Extend approach to detailed patterns, such as birdcalls and other elements of the biophony
- Establish reliability of using MP3 spectral frame data
- Classification algorithm development

Acknowledgements

- **Co-author Joe Studniarz (undergraduate research assistant at Montana State Univ.)**
- **NPS Natural Sounds and Night Skies Program, Ft. Collins, CO**
  - Kurt Fristrup, Emma Lynch, Damon Joyce
- Montana Institute on Ecosystems
  - Cathy Whitlock and Ric Hauer, Directors

Soundscape collaboration for science, management, and public outreach at a national historic site

Robert C. Maher

Montana State University – Bozeman

Christine Ford

Grant-Kohrs Ranch NHS, MT

162nd Meeting of the Acoustical Society of America – San Diego – November 2011

Outline

- Introduction
- Collaboration in soundscape research
  - Scientists and Engineers
  - Site managers
  - Public outreach
- Project Example: Grant-Kohrs Ranch National Historic Site, Deer Lodge, MT
- Prospects for future work
- Conclusion

Introduction

- Scientists and engineers can learn a great deal by working with park managers and public outreach professionals.
- What is needed:
  - Common language
  - Understanding what is possible and what is not using current technology
  - Mutual learning and respect

Example: Grant-Kohrs Ranch National Historic Site

- Deer Lodge, Montana
- A working cattle ranch commemorating the heritage of American cowboys, stock growers, and cattle operations during the late 19th and early 20th centuries.
- Congress: maintain the site as a working ranch.
- Cultural soundscape is essential: all the sights, sounds, and sensations associated with ranching.

MONTANA STATE UNIVERSITY | College of ENGINEERING

Mountains & Minds

Getting the ball rolling

- In 2008, Grant-Kohrs NHS managers were curious (and concerned) about soundscape
- BUT…no on-site expertise, and only minimal funding available
- NHS contacts National Park Service Natural Sounds Office
- Montana State University contacted: able to work within geographic and funding constraints via Rocky Mountain Cooperative Ecosystem Study Unit (RM-CESU)

The Collaboration

- National Historic Site: management and interpretation for the public
- National Park Service professionals: expertise and equipment
- Montana State University: audio signal processing expertise—plus education, training, and research opportunity

Long-Term Collection


March 17, 2009


September 5, 2009

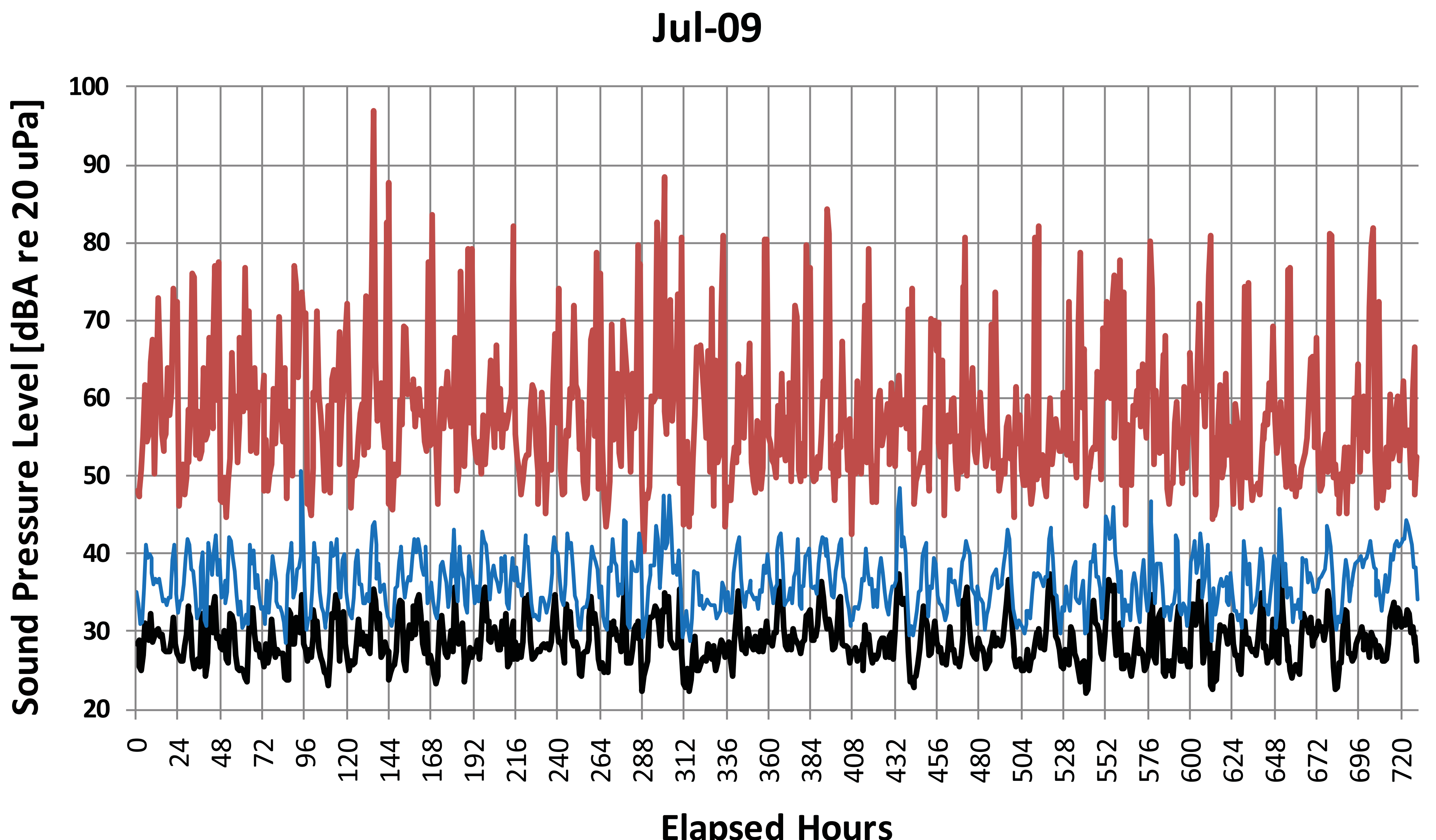

June 22, 2009


December 12, 2009

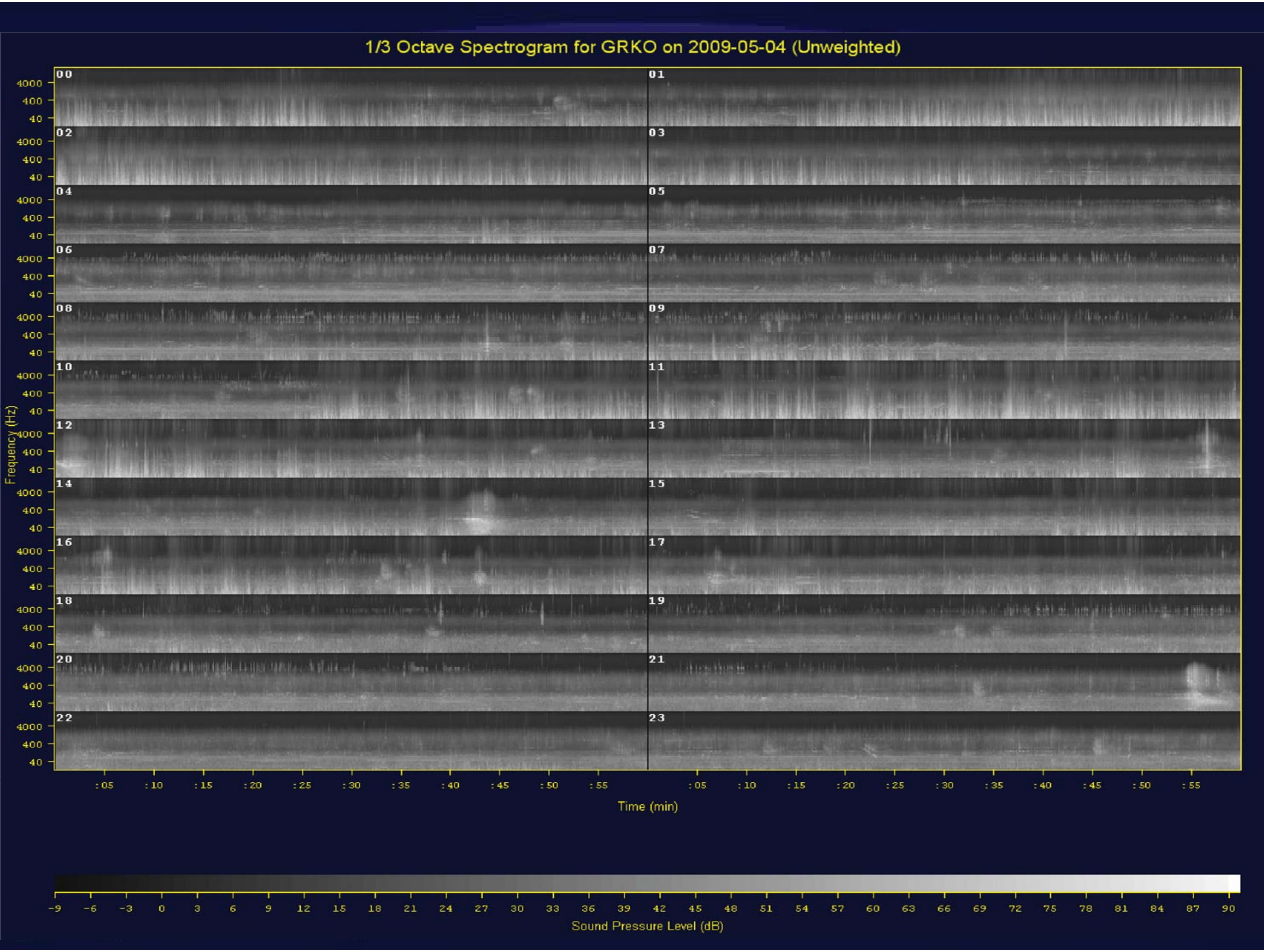
Project Outcomes

- Audio recording lasting 365 days
  - 8,760 hours (lots of MP3 files)
  - 1 second 1/3rd octave sound levels
  - 10 second windspeed
- Long segments of natural quiet with sections of recognizable biophony, geophony, and anthrophony

SPL Graphs



Jul-09
Sound Pressure Level [dBA re 20 uPa]
100
90
80
70
60
50
40
30
20
0
24
48
72
96
120
144
168
192
216
240
264
288
312
336
360
384
408
432
456
480
504
528
552
576
600
624
648
672
696
720
Elapsed Hours


1/3 Octave Spectrogram for GRKO on 2009-05-04 (Unweighted)
00
01
02
03
04
05
06
07
08
09
10
11
12
13
14
15
16
17
18
19
20
21
22
23
Frequency (Hz)
4000
400
40
:05 :10 :15 :20 :25 :30 :35 :40 :45 :50 :55
Time (min)
-9 -6 -3 0 3 6 9 12 15 18 21 24 27 30 33 36 39 42 45 48 51 54 57 60 63 66 69 72 75 78 81 84 87 90
Sound Pressure Level (dB)

Prospects for future work

- Treasure trove or a data explosion?
- Finding the needle in the haystack, or not seeing the forest for the trees?

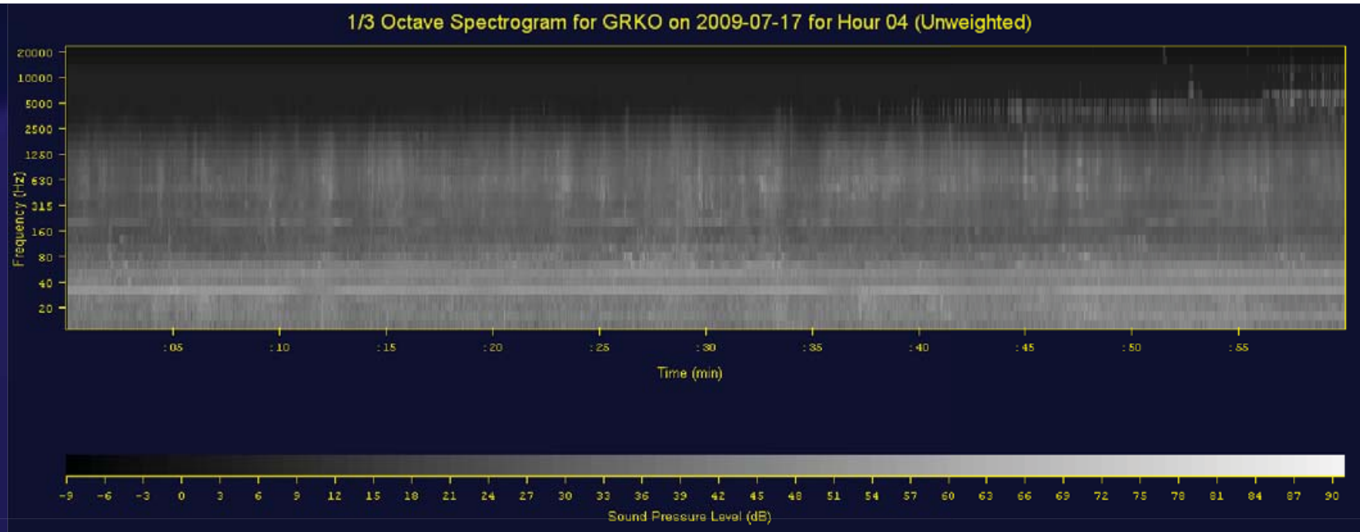
- Matching prospects and expectations for systematic evaluation.

Conclusion

- It's fun to work with smart people who are passionate about what they do. The National Park Service is full of these sorts of people!
- Challenges and resources are naturally interpreted differently by each group
- Collaboration is professionally rewarding, even if not economically rewarding

Acknowledgements

- Grant-Kohrs Ranch NHS, Deer Lodge, MT
- NPS Natural Sound and Night Sky Program, Ft. Collins, CO
- Rocky Mountain Cooperative Ecosystem Studies Unit (RM-CESU), Missoula, MT



Automated analysis and interpretation of long-term soundscape audio recordings

Robert C. Maher

Electrical and Computer Engineering

Montana State University - Bozeman

MONTANA STATE UNIVERSITY | College of ENGINEERING

Mountains & Minds

Outline

- Introduction
- Long-term acoustical acquisition
  - Soundscape studies (thousands and thousands of hours)
  - Environmental monitoring
- How to present soundscape information?
  - Spectrograms
  - Aural snapshots
  - Time-lapse aural presentation
- Conclusion

Introduction

- Long-term soundscape studies are now feasible and desirable
- Interpretation and presentation is difficult due to extreme length of the data
- What is needed:
  - Automated analysis tools
  - Useful comparison metrics
  - Meaningful presentation techniques

Example: Grant-Kohrs Ranch National Historic Site

- Deer Lodge, Montana
- A working cattle ranch commemorating the heritage of American cowboys, stock growers, and cattle operations during the 19th and 20th centuries.
- Congress: established in 1977 to maintain the site as a working ranch.
- Cultural soundscape is essential: all the sights, sounds, and sensations associated with ranching.

MONTANA STATE UNIVERSITY | College of ENGINEERING

Mountains & Minds

MONTANA STATE UNIVERSITY | College of ENGINEERING

Mountains & Minds

Long-Term Collection

Project Presentation Challenges

- Audio recording lasting 365 days = 8,760 hours (525,600 minutes)
- Long segments of natural quiet with sections of recognizable biophony, geophony, and anthrophony
- Visitors to a web site or visitor center spend only a few minutes: can we compress meaningfully by 1/200,000 ?

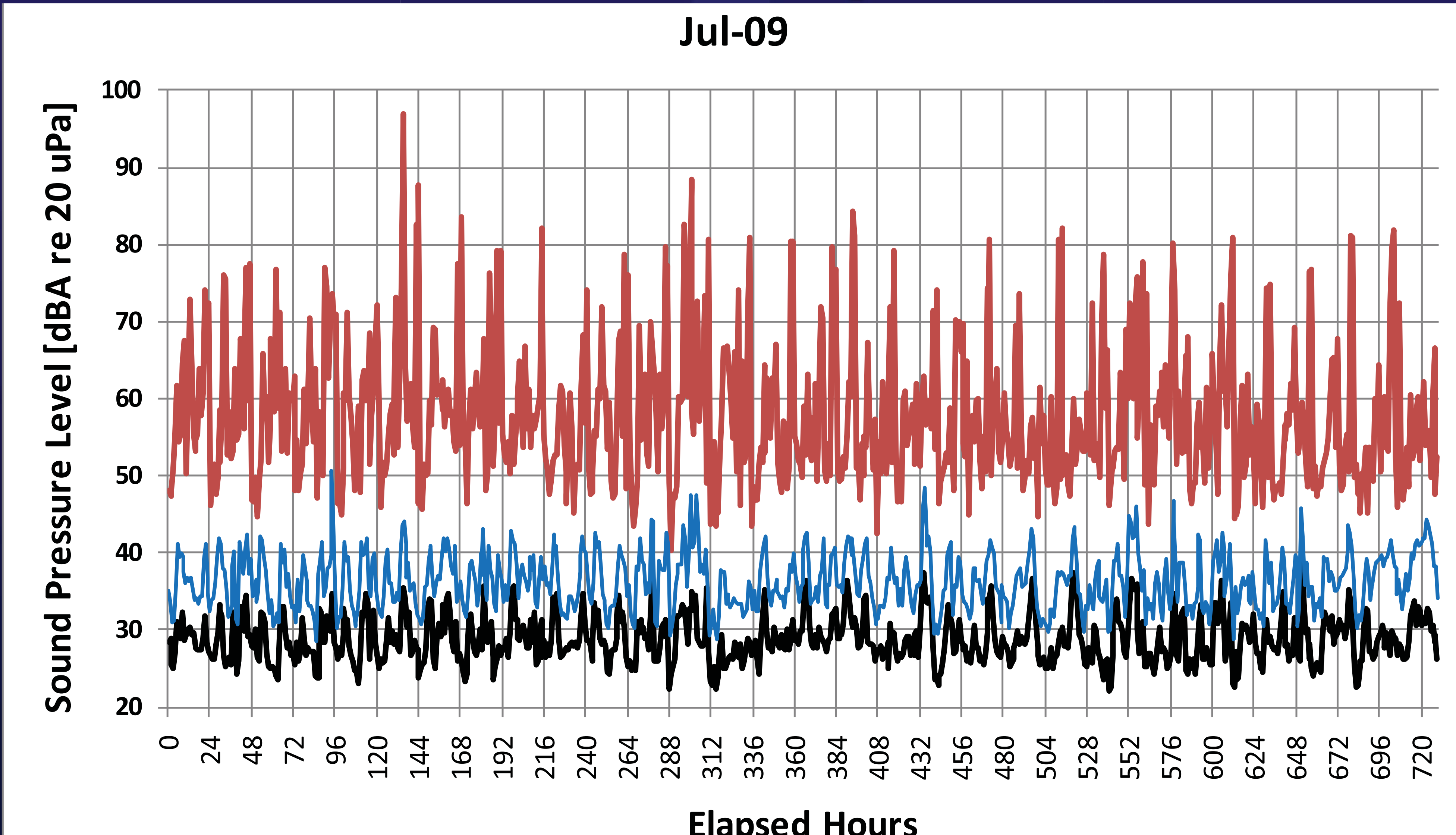
Some options

- Automated SPL min/max/average graphs
- Spectrographic displays
- Audio samples of “highlights”
- Time-lapse aural display

SPL Graphs

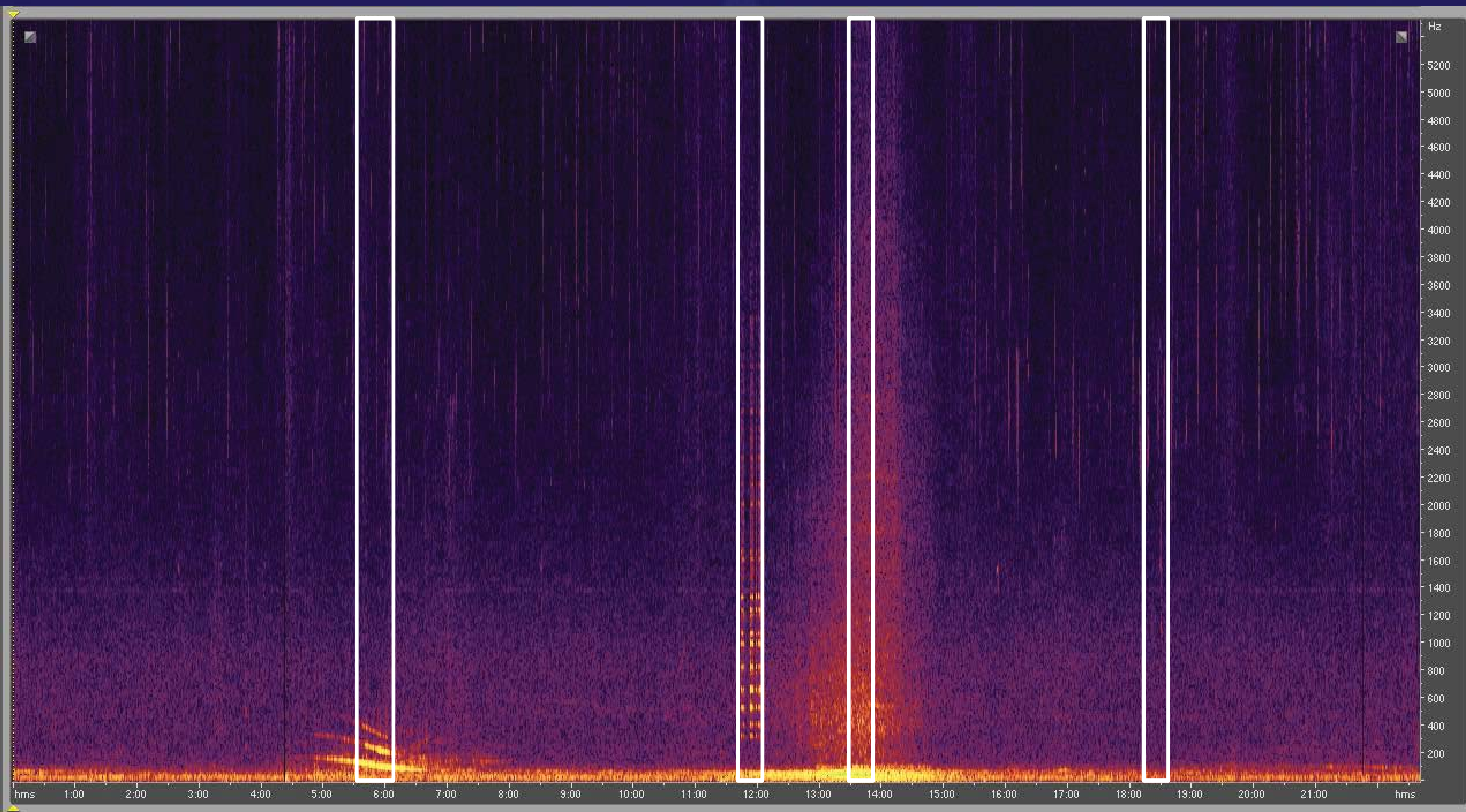
- Presents information on maximum, minimum, and average sound levels
- Relatively simple to produce
- Interpretation still required
- Little information in general about sound sources and distributions

SPL Graphs


Jul-09
Sound Pressure Level [dBA re 20 uPa]
100
90
80
70
60
50
40
30
20
0 24 48 72 96 120 144 168 192 216 240 264 288 312 336 360 384 408 432 456 480 504 528 552 576 600 624 648 672 696 720
Elapsed Hours

Spectrographic Display

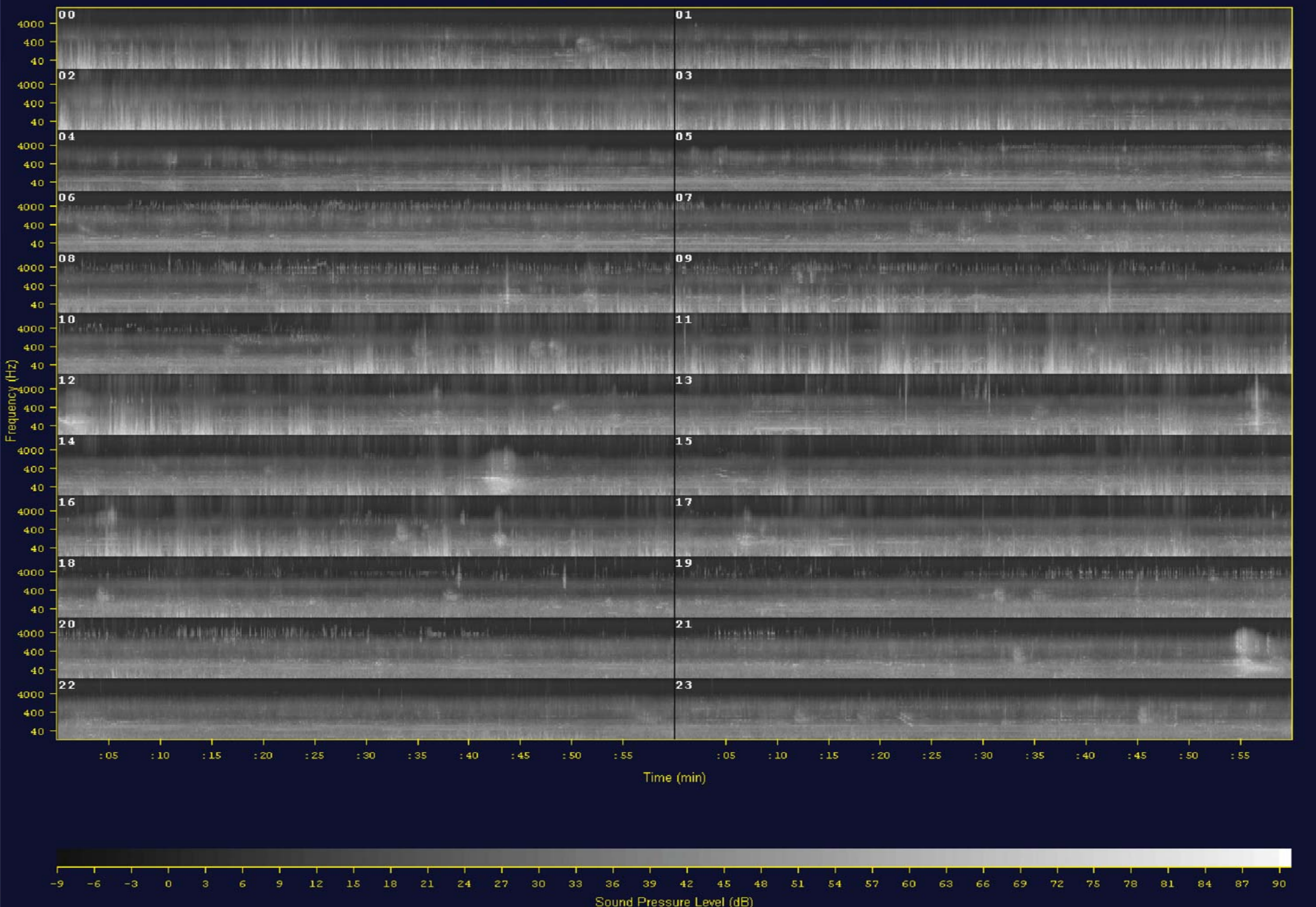
- Conveys time-frequency-energy distribution
- Condenses a lot of information into a compact form
- May be confusing to the public unless explained
- Works best if audio playback allowed (point and click)



MONTANA STATE UNIVERSITY | College of ENGINEERING

Mountains & Minds

1/3 Octave Spectrogram for GRKO on 2009-05-04 (Unweighted)



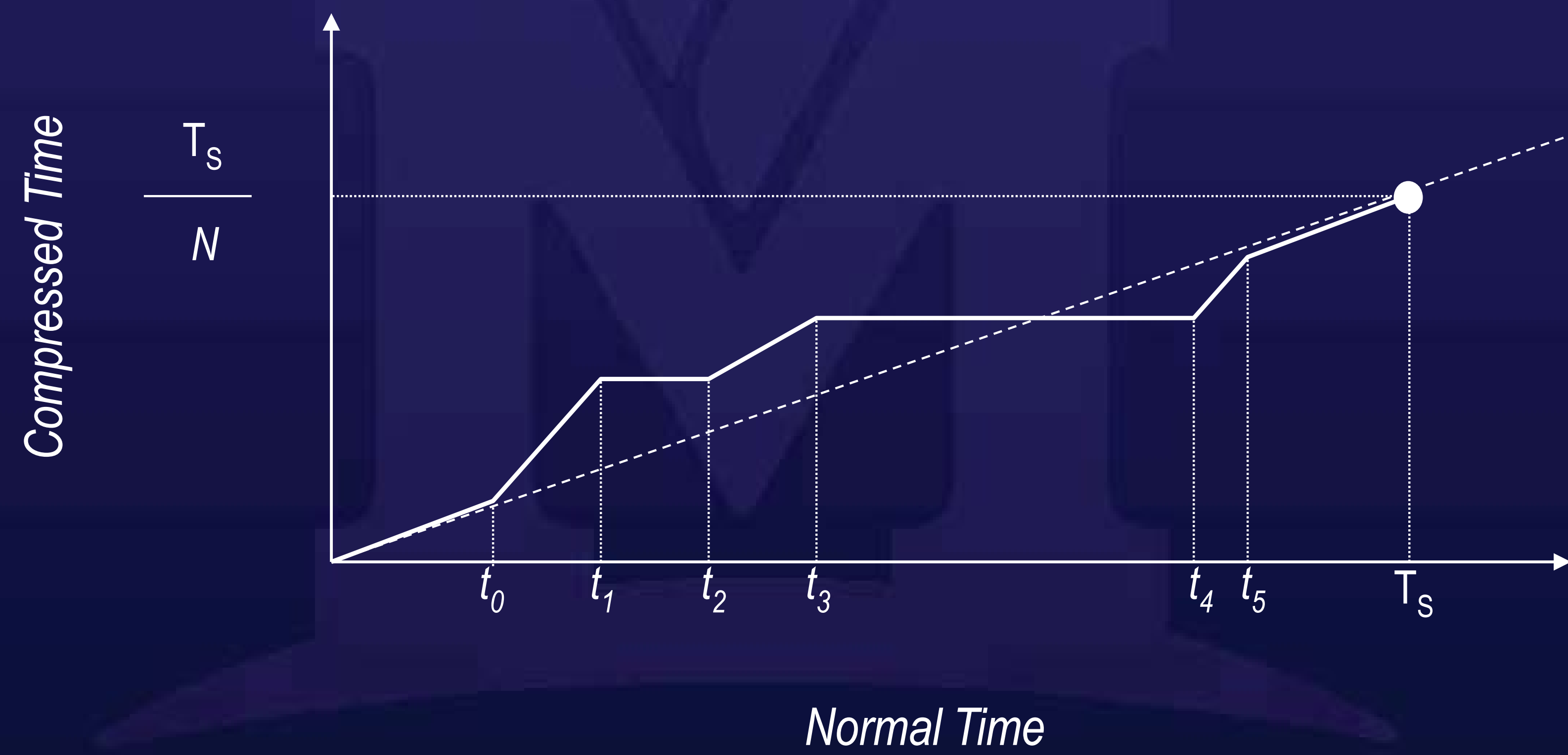
Audio highlights excerpts

- Identify and extract “interesting” sound examples
- The visitor can quickly sample the range of sounds and sound textures
- Generally requires considerable audition and manual preparation
- May give a non-representative indication of the actual sound texture

Time-lapse aural display

- Goal: represent the aural sound texture for many minutes of real time audio with only a few seconds of seamless excerpts
  - Aural equivalent of time-lapse photography
- Challenge: defining and capturing sound texture in an aurally meaningful manner
  - Simple block-downsampling may not capture sonic *texture* effectively

Non-uniform time warp concept

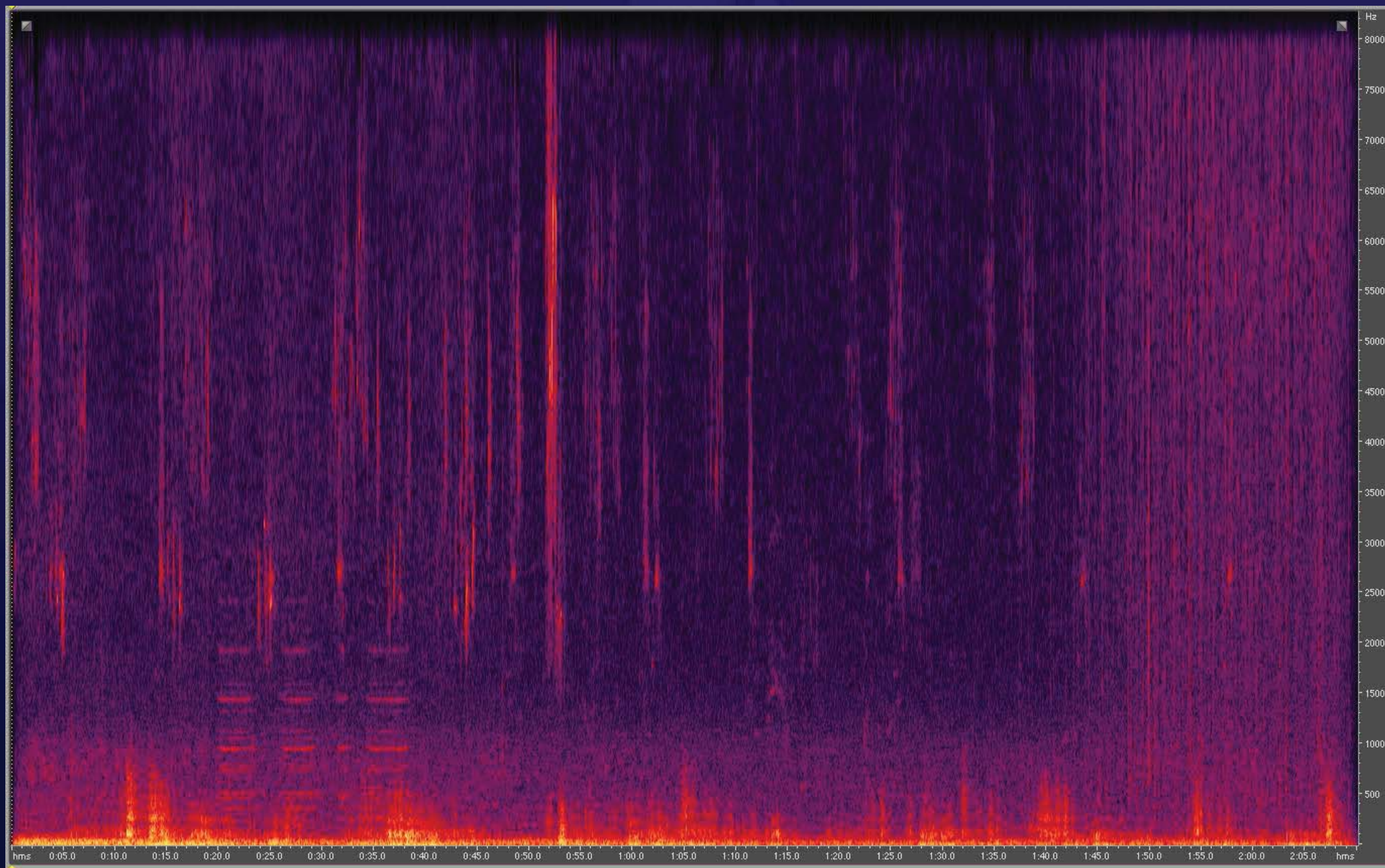


Approach for time-lapse aural display

- Create a spectral transition map: identify textural boundaries in the audio
- Determine available segments based on compression factor N
- Assign segments to the transitions in order of priority
- Segment the audio and concatenate with overlap-add

Example Spectrogram

8kHz

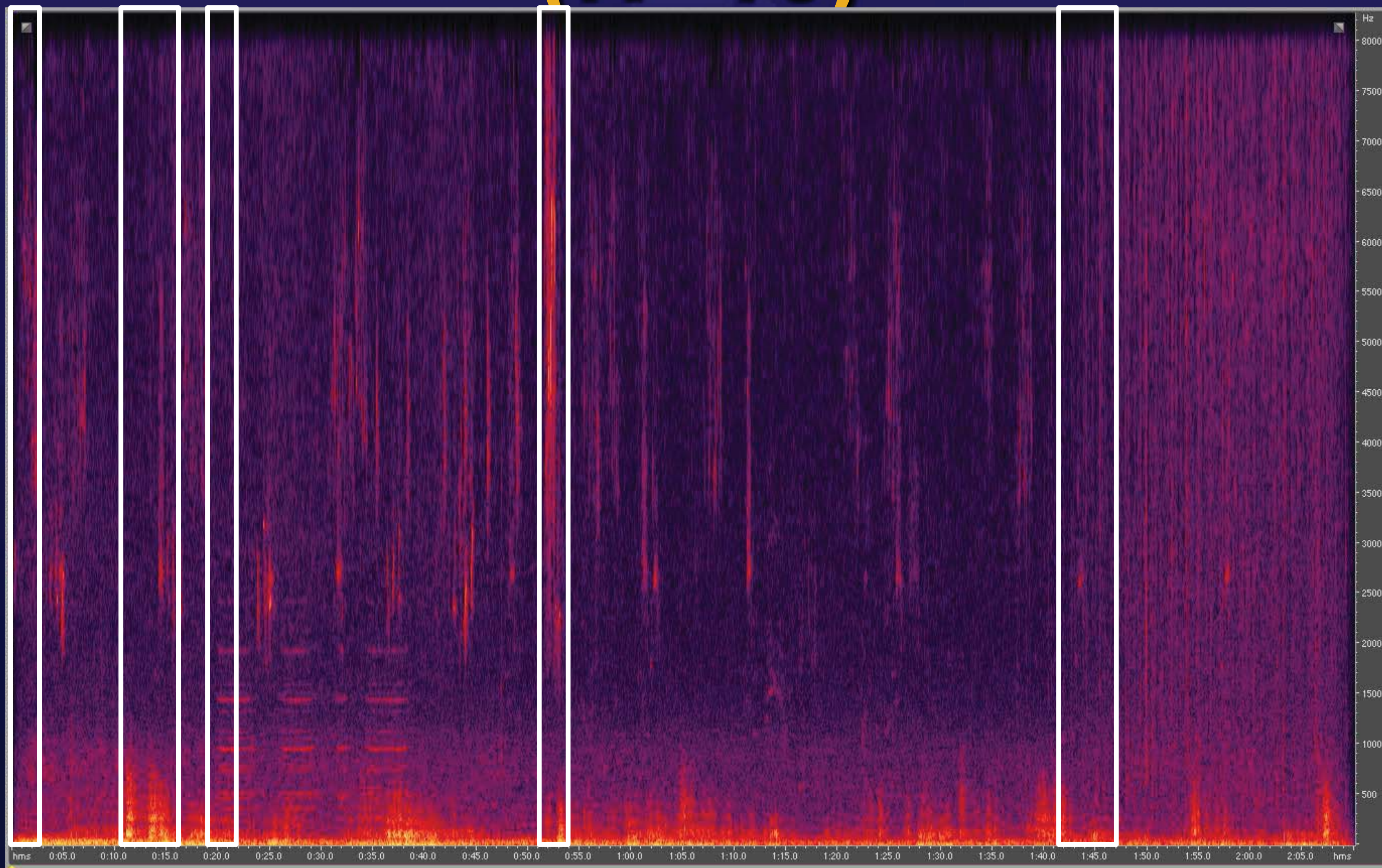


0

2' 10"

Example Transition Map (N=10)

8kHz

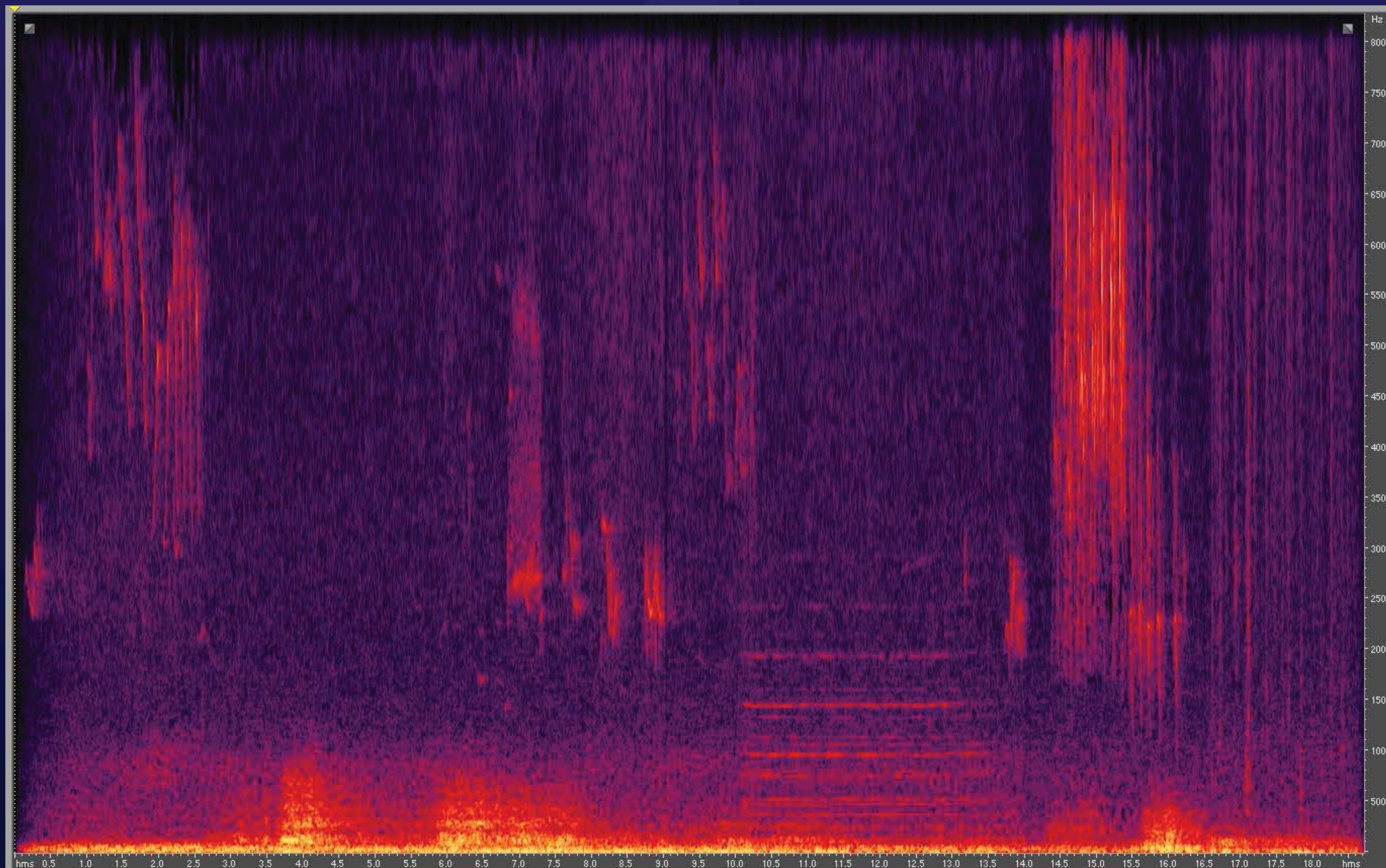


0

2' 10"

Reconstructed Signal (N=10)

8kHz



0

18”

MONTANA STATE UNIVERSITY | College of ENGINEERING

Mountains & Minds

Conclusion

- Long-term acoustical acquisition requires automated analysis and distillation tools
- Presenting days/weeks/months of audio is challenging
- Extreme time-scale compression is necessary for many applications
- Ongoing effort is needed in pattern detection and pattern matching

Acknowledgements

- **Grant-Kohrs Ranch NHS, Deer Lodge, MT**
  - Christine Ford, Integrated Resources Program Manager
- **NPS Natural Sounds Program, Ft. Collins, CO**
  - Kurt Fristrup, Emma Lynch, Damon Joyce
- **Rocky Mountain Cooperative Ecosystem Studies Unit (RM-CESU), Missoula, MT**
  - Kathy Tonnessen, Natural Resources Research Coordinator
  - Lisa Gerloff, Executive Coordinator

Sound Examples

- March 18, 2009 9:34PM MDT (45”)
- April 15, 2009 6:13AM MDT (before dawn)(1’)
- May 1, 2009 11:22AM MDT (5’)
- May 4, 2009 6:23AM MDT (after dawn) (2.5’)
- July 6, 2009 ~noon (6.5’)
- Dec 30, 2009 9:30PM MST (2.5’)
- July 6, 2009 ~1:30PM MDT (2’)

- http://ece.montana.edu/rmaher/audio_monitor/grko.htm

Audio Engineering Society

Convention Paper

Presented at the 129th Convention
2010 November 4–7 San Francisco, CA, USA

Maintaining sonic texture with time scale compression by a factor of 100 or more

Robert C. Maher

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

ABSTRACT

Time lapse photography is a common technique to present a slowly evolving visual scene with an artificially rapid temporal scale. Events in the scene that unfold over minutes, hours, or days in real time can be viewed in a shorter video clip. Audio time scaling by a major compression factor can be considered the aural equivalent of time lapse video, but obtaining meaningful time-compressed audio requires interesting practical and conceptual challenges in order to retain the original sonic texture. This paper reviews a variety of existing techniques for compressing 24 hours of audio into just a few minutes of representative "time lapse" audio, and explores several useful modifications and challenges.

1. INTRODUCTION

The increasing availability of audio recording devices capable of capturing hours, days, or weeks of continuous audio data has opened many new possibilities for sonic analysis, especially in understanding the diurnal and seasonal changes in natural sound environments [1, 2]. Although digital storage of such recordings is now widely available at a steadily decreasing cost, the notion that a human listener would have the time—and aural concentration and stamina—to listen to weeks of continuous audio seems highly unlikely. Thus, such long-term recordings will require automated processing and effective time-scale compression in order to be used and studied.

Typical time scaling techniques fall into two categories: time domain techniques such as the method of Roucos and Wilgus [3], and frequency domain techniques such as the phase vocodor [4]. For time scale compression (i.e., output_length = input_length/Factor) up to a factor of two or so, techniques that separate the temporal envelope from the underlying periodic excitation signal and scale them separately work quite well for speech and other quasi-periodic signals.

However, for time compression factors of 20, 50, 100, or more, the vast majority of the original audio temporal information must be eliminated. While this major time compression will likely destroy the intelligibility of a continuous speech recording, the more common situation in recordings of natural sound environments is a plethora of overlapping sounds with distinct patterns in frequency and time that create an ensemble typically referred to as a *sonic texture*. For example, a recording from a wilderness area or a backcountry location in a national park might contain subtle overlapping *background* sounds such as running water, bird calls, wind, insects, amphibians, high-altitude passenger jets, etc., with an occasional *foreground* sound like footsteps or human conversation. These sounds would come and go throughout the hours of the recording.

The desirable features of a time-compressed version of the recording would be to retain the sequencing and relative prevalence of the various sound sources in the audio time lapse version. An individual listening to the time-compressed version should be able to hear the individual sound sources and judge the overall sonic texture as being consistent with the original real time recording.

This concept is depicted in Figure 1. The normal time axis is mapped linearly to the compressed time axis with slope $1/N$ (dashed line), where N is the time compression factor. However, rather than applying uniform time compression to the entire segment, the algorithm identifies portions of the input signal containing foreground sounds for reduced time compression (segments t_0-t_1, t_2-t_3, and t_4-t_5), portions with consistent background texture (segments 0-t_0 and t_5-T_S), and portions consistent with the adjacent textures and therefore redundant (segments t_1-t_2 and t_3-t_4). This results in a time warping function that is still monotonically increasing, but no longer a simple straight line mapping.

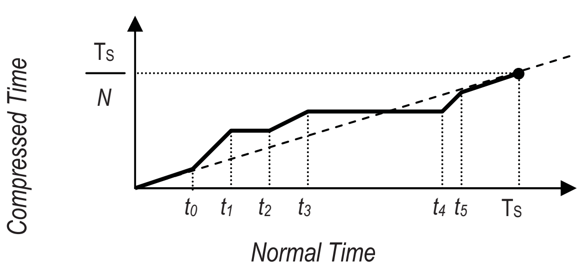


Figure 1: Time-compression by an overall factor N, with non-uniform warping to retain sound events.

Although the time warping function could conceivably be used to drive a conventional time-scale modification algorithm, the extreme time compression envisaged in this project (e.g., 60:1 or 100:1) would result in completely unnaturally compressed duration of individual sound events. Instead, a more holistic view and algorithm is needed.

There have been a variety of published papers in the area of major time-compression for audio. In addition to the techniques designed for modest compression factors of speech and music, there have been several notable efforts specifically dealing with sound textures. The contemporary composer R. Luke DuBois has coined the term "time lapse *phonography*" to describe his creative compositions involving sampled music segments [5]. Crockett and others have used the concepts of computational auditory scene analysis (CASA) to identify distinct sonic elements in a recording and to treat them individually [6]. In a complementary area, Frojd and Horner recently used the concepts of granular synthesis to create plausible sonic textures of auditory scenes such as applause and crowd noise, thereby synthesizing audio textures of arbitrary length [7].

One simple time-domain approach for extreme time scale compression is to do *block downsampling*. Short segments of the input signal are taken from widely spaced locations such that we select one segment out of N to produce the N:1 compression factor. Each short segment is windowed and then overlap-added to produce the time-compressed output sequence. The procedure begins by identifying a suitable block length, T, typically between a few hundred up to a few thousand milliseconds, which is chosen to be long enough to contain recognizable sounds, but short enough to allow suitable quality for the time lapse effect. The block downsampling method is shown in Figure 2.

The block downsampling approach is simple and straightforward to implement, but has the drawback that the arbitrary block segmentation does not take into account the details of the underlying signal itself, and therefore the sonic integrity of the output signal may be degraded. Nevertheless, it seems plausible that this drawback could be reduced by analyzing the input signal to select the location of the downsampling blocks to capture the individually recognizable foreground sounds, while still retaining the background sonic texture. This observation is the starting point for the work described in this paper.

The proposed method includes an automatic analysis/statistics stage and a synthesis/reconstruction stage. The automatic analysis stage determines a reasonably coarse description of the background sonic texture, and a reasonably fine segmentation of the foreground sonic events, which are relatively strong and presumably recognizable individual sounds.

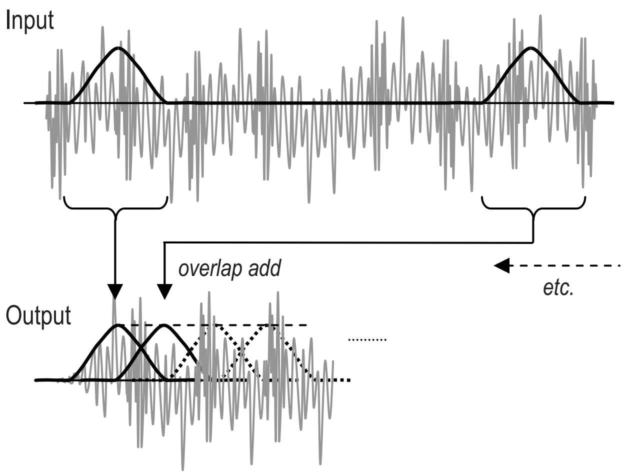


Figure 2: Simple time-domain block downsampling procedure.

2. SONIC TEXTURE: STATISTICS AND ANALYSIS

The goal of this work is to allow extreme time compression of audio recordings while maintaining sonic texture and distinctive foreground sounds. If we consider an example in which we seek a 60:1 compression, that is, keeping the equivalent of one minute out of every hour, we must blend or eliminate 98% of the temporal information while retaining the useful essence of the audio material. Thus, it is essential to have a practical and objective measure of the rather subjective concept of sonic texture.

For this investigation we start with a one second time interval and a 1/3rd octave frequency interval. Neither choice is necessarily optimum psychoacoustically, but the one second time interval and 1/3rd octave frequency interval are found to be sufficiently fine to capture the time-variant spectral character of the signal. Each one-second frame generates thirty 1/3rd octave band levels (20 Hz to 20 kHz range). The texture is defined by the time-variant fluctuation in the 1/3rd octave band levels.

The textural criterion that has been most useful has been to identify via automatic software analysis the frame-to-frame level differences in one or more bands that exceed a threshold, such as 1.5 dB. For each frame, the process looks at the next frame and at the subsequent five to ten frames to determine any repetitive fluctuations. The result is a map of the textural transitions.

As an example, consider the 1/3rd octave spectrogram of a 60 minute signal shown in Figure 3a. The primary background sonic texture of this signal includes a sustained low frequency component and a fluctuating mid-frequency component. There are also a few distinct events that are relatively short in duration, and some high frequency detail.

Applying the analysis algorithm with the frame-to-frame level difference criterion and a ten-frame median smoothing, the resulting background texture transition map is shown in Figure 3b. Note that the map indicates a few minutes of spectral continuity, punctuated by the distinct events and transitions. The goal is to retain the time segments with a high number of sonic texture transitions, at the expense of the segments with lower activity.

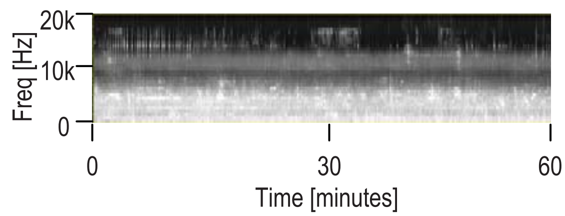


Figure 3a: Example 1/3rd octave spectrogram for 60 minute example signal.

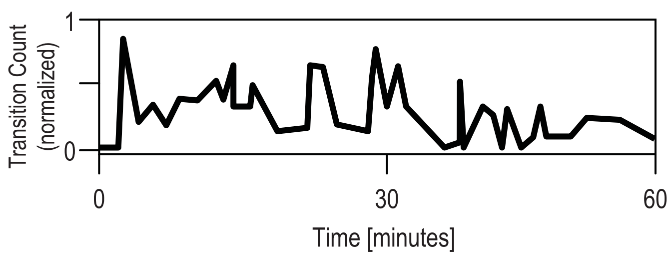


Figure 3b: Texture transition map analyzed from the spectrogram of Figure 3a.

Identifying the foreground sonic events is accomplished by the automatic software by looking longitudinally in groups of 1/3rd octave bands. If a particular band group shows an increase in level over one or two frames, the frame is flagged for event consideration. Applying this analysis to the example in Figure 3a produces the list of candidate foreground events shown in Figure 3c.

Once the entire length of the input audio file is processed, the rate of occurrence of the textural transitions and the number of foreground event flags is calculated. The idea is to look at the rate of occurrence of texture blocks and the timing of the event frames, and to attempt to maintain this rate after time compression.

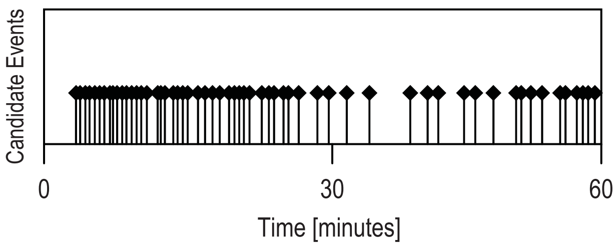


Figure 3c: Distinct sonic event map analyzed from the spectrogram of Figure 3a.

Reviewing the example of Figure 3a-c, the analysis indicates more than 50 distinct sonic events and at least a dozen texture transitions within the 60 minute signal segment. If the goal is a 60:1 time compression, i.e., on average each minute of the original recording becomes one second in the compressed version, the processing will focus on retaining the distinct events and the transitions, although it is clear that the duration of each event will need to be curtailed to fit everything into the time-compressed output.

3. LONG-TERM TEXTURE EXAMPLE

A more lengthy and complicated example will help illustrate the difficulties of the analysis procedure. A 24 hour segment of natural sound recorded at the Grant-Kohrs Ranch National Historic Site (Deer Lodge, Montana, USA) is shown in Figure 4, displayed as a 1/3rd octave spectrogram. Segment “00” in the top left of Figure 4 corresponds to the midnight – 1:00AM hour, segment “01” in the upper right corner corresponds to the 1:00AM – 2:00AM hour, and so forth. The brightness of the spectrogram details indicates the sound pressure level at that particular time and frequency range.

In order to allow a reasonable duration for audition, the 24 hours of audio ideally would be compressed into 12 minutes or so (120:1). This is reasonable for many of the hours in which the sonic texture is consistent and subtle, such as hours 00 through 05, but is difficult to handle for more lengthy sonic events, such as a passenger jet flyover during hour 12, and train whistles and track noise during hour 21 and during hour 23. Additional research is needed to determine the best way to time compress events like the flyover and the train sounds that last many minutes in normal time, but cannot occupy more than a few seconds in the compressed-time regime. This remains an open question.

4. SEGMENTATION AND ASSEMBLY

With 98% or more of the temporal audio frames discarded by extreme time scale compression, the segmentation and reassembly requires both elimination of redundant frames and seamless mixing and concatenation of the frames that are essential to maintain the sonic texture. In many examples the challenge is that the analysis detects more sonic events and texture transitions than can be accommodated in the time compressed output. Some priority ranking must be deduced or inferred to decide which events and transitions will be left out of the compressed result.

The current solution to this issue uses the windowed overlap-add to recreate the time compressed output signal. The assembly and reconstruction occurs in the time domain, with the distinct sonic event map guiding the selection of high-priority frames, and the texture transition map guiding the selection of background frames to fill in the required duration between the foreground sonic events.

5. CONCLUSION

Extreme time scale compression of audio material for time-lapse purposes has no naturally occurring counterpart, so establishing valid metrics for assessing quality and fidelity is largely a subjective process. Future work will be needed to quantify the degree to which the time compressed output signal retains the texture and sonic character of the input signal. In the meantime, the basic compression strategy and goals described in this paper have been found to be useful and practical.

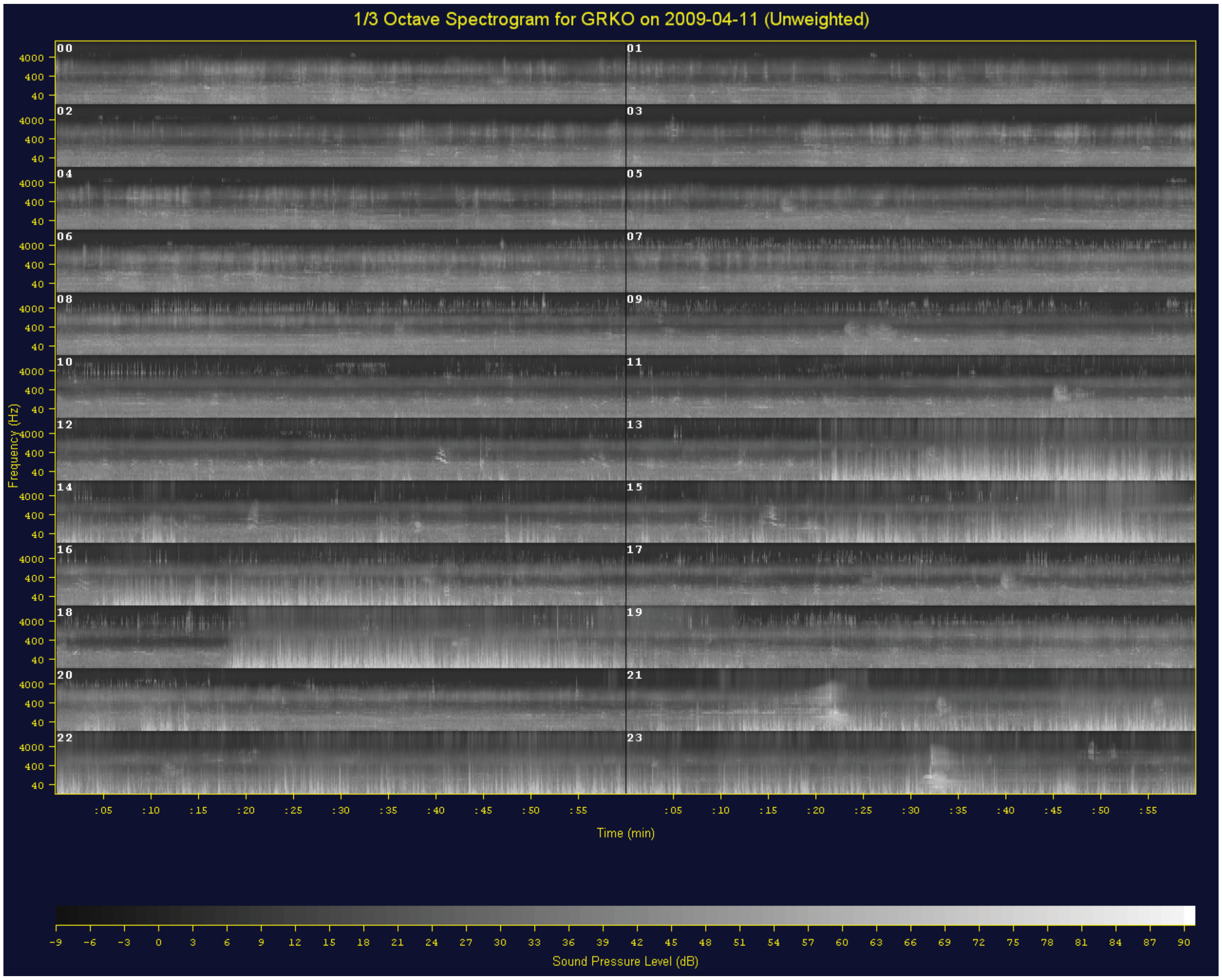


Figure 4: 24-hour spectrographic summary of the Grant-Kohrs Ranch monitor site on 2009 April 11.

6. ACKNOWLEDGEMENTS

This work was supported in part by the U.S. National Park Service through a contract with the Rocky Mountain Cooperative Ecosystem Studies Unit (RM-CESU). The author gratefully acknowledges the assistance of Chris Ford, Grant-Kohrs Ranch National Historic Site (Deer Lodge, Montana, USA), and Damon Joyce, National Park Service Natural Sounds Program (Fort Collins, Colorado, USA).

7. REFERENCES

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

[2] Maher, R., "Acoustics of national parks and historic sites: the 8,760 hour MP3 file," Proc. AES 127th Convention, preprint 7893, Oct 2009.

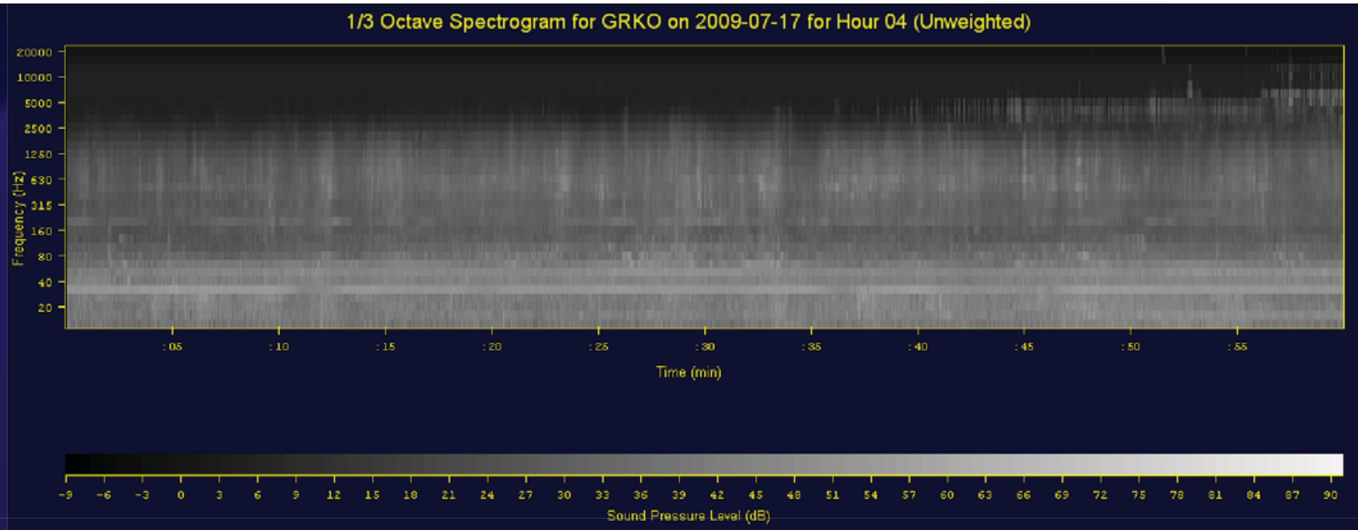
[3] Roucos S. and Wilgus A.M., "High quality time-scale modification for speech", Proc. IEEE ICASSP, pp. 493-496, March 1985.

[4] Laroche, J., and Dolson, M., "Improved phase vocoder", IEEE Trans. Speech and Audio Processing, vol 7, no. 3, pp. 323 –332, May 1999.

[5] DuBois, R. Luke, http://lukedubois.com/

[6] Crockett, B., "High quality multi-channel time-scaling and pitch-shifting using auditory scene analysis, Proc. AES 115th Convention, preprint 5948, Oct 2003.

[7] Frojd, M., and Horner, A., "Sound texture synthesis using an overlap–add/granular synthesis approach," J. Audio Eng. Soc., vol. 57, no. 1/2, pp. 29-37, Jan/Feb 2009.



Maintaining sonic texture with time scale compression by a factor of 100 or more

Robert C. Maher

Electrical and Computer Engineering

Montana State University - Bozeman

MONTANA STATE UNIVERSITY | College of ENGINEERING

Mountains & Minds

Outline

- Introduction
- Long-term acoustical acquisition
  - Soundscape studies (thousands and thousands of hours)
  - Environmental monitoring
- <u>Problem</u>: how to present the data
- Sonic texture: statistics and analysis
  - 1/3rd octave band analysis
  - Frame-to-frame level changes
  - Sonic event map
  - Examples
- Conclusion

Introduction

Sound *texture*

- A loosely-defined term intended to express
  - Low-level *background* sounds
  - Occasional distinctive *foreground* sounds
  - Reverberation characteristics
  - Spectro-temporal trends
- There is an assumption of self-similarity on various time scales

Introduction (cont.)

Why is there any interest in extreme time compression?

- Long-term soundscape studies
- National Park Service sound management
  - Example: Grant-Kohrs Ranch (GRKO) National Historic Site
- Similarity to time-lapse photography: fast depiction of a slow temporal scene

National Park Service Act (1916)

- *"...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."*

NPS Management Policies 2006

National Park Service *Management Policies 2006* include natural and cultural sound resources within park units.

- 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

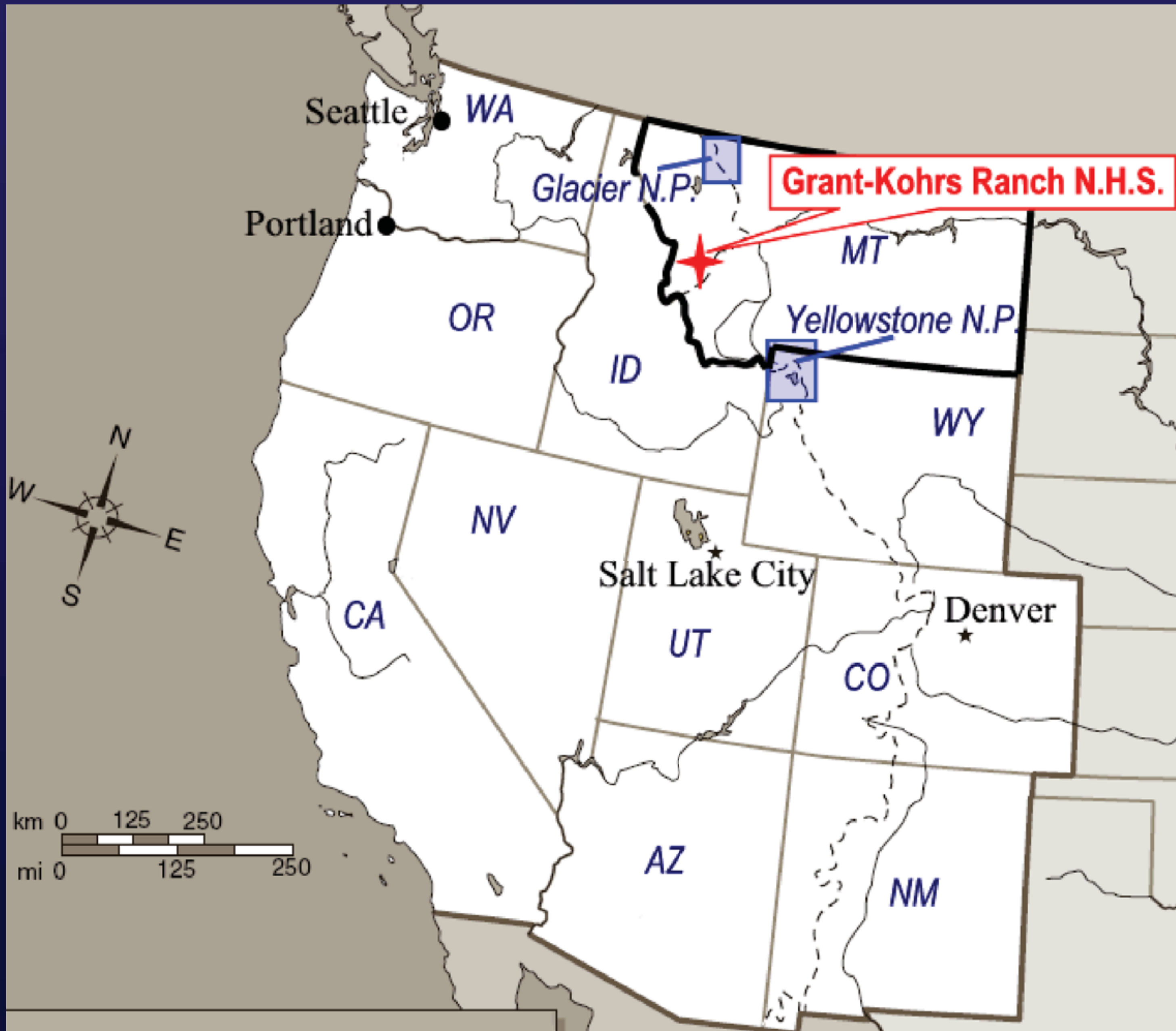
Soundscape Regulatory Context

- 1872 Yellowstone National Park Act
- 1916 National Park Service (NPS) Organic Act
- 1949 Executive Order 10092 (Boundary Waters no-fly zone)
- 1964 Wilderness Act
- 1969 National Environmental Policy Act
- 1972 Noise Control Act
- 1987 National Parks Overflights Act (NPOA)
- 1988 Special Federal Aviation Regulation (SFAR) 50-2 (GRCA)
- 2000 National Parks Air Tour Management Act
- 2000 NPS Director's Order #47 (soundscape preservation)
- 2002 Winter Use Plan (Yellowstone)
- 2006 NPS Management Policies (soundscapes)

- Miller, Nicholas, P., "US National Parks and management of park soundscapes: a review," Applied Acoustics, vol. 69(2), pp. 77-92, February 2008
- R.C. Maher, J. Gregoire, and Z. Chen, "Acoustical monitoring research for national parks and wilderness areas," Preprint 6609, Proc. 119th Audio Engineering Society Convention, New York, NY, October 2005.

Grant-Kohrs Ranch National Historic Site (1977)

- Deer Lodge, Montana
- A working cattle ranch commemorating the heritage of American cowboys, stock growers, and cattle operations during the 19th and 20th centuries.
- Congress: maintain the site as a working ranch.
- Cultural soundscape is essential: all the sights, sounds, and sensations associated with ranching.


Grant-Kohrs Ranch N.H.S.
Seattle
WA
Glacier N.P.
Portland
MT
Yellowstone N.P.
OR
ID
WY
N
W
E
S
NV
Salt Lake City
CA
Denver
UT
CO
km 0
125
250
mi 0
125
250
AZ
NM

MONTANA STATE UNIVERSITY | College of ENGINEERING

Long-Term Sound Collection

Soundscape

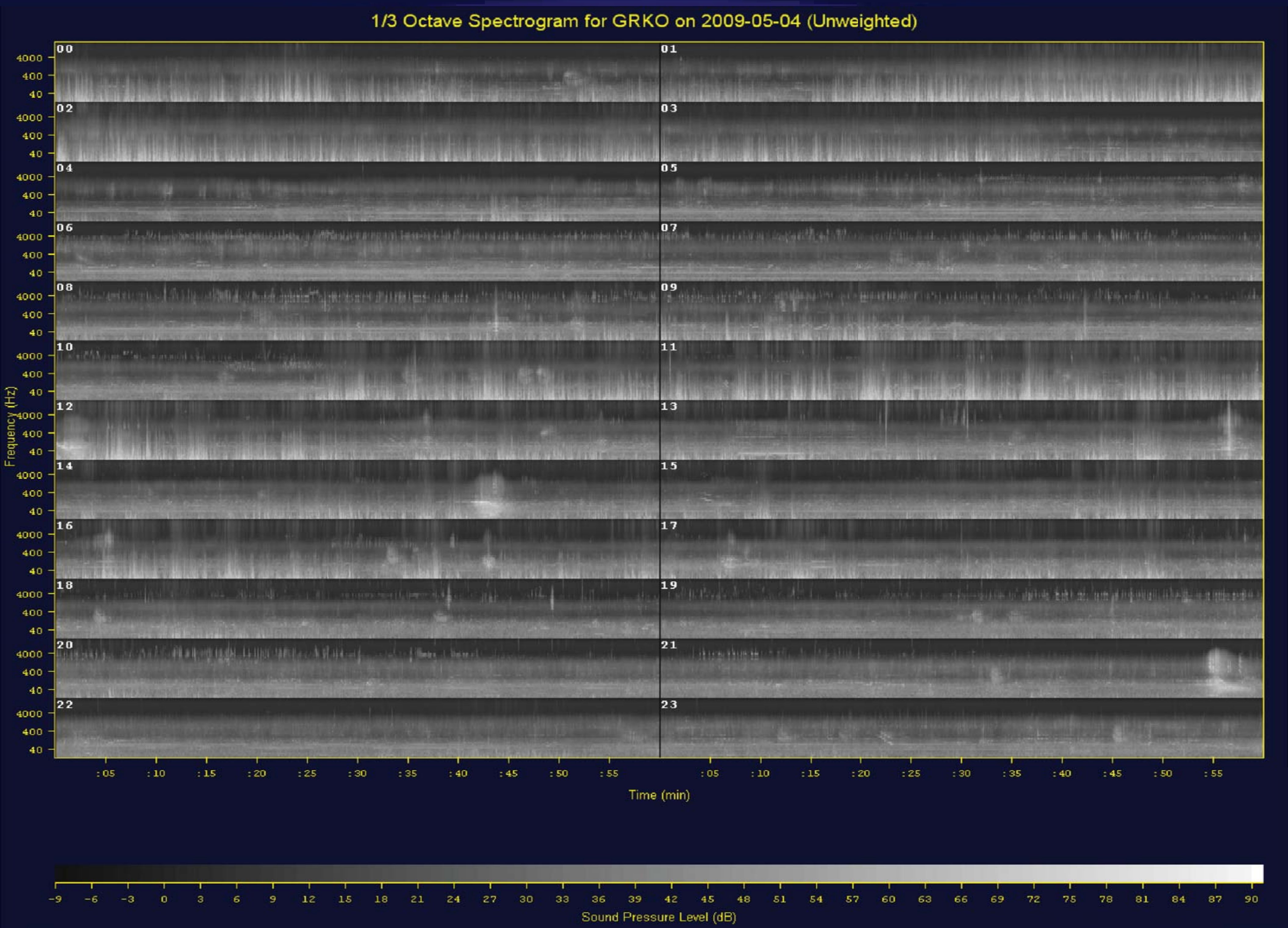
Three Sonic Components (*Krause*):

- ***Biophony*** -- animal and biological sounds
- ***Geophony*** -- geological, hydrological, and meteorological sounds
- ***Anthrophony*** -- sounds caused by humans and human activity

NPS GRKO Project Results

- For 365 days:
  - 1 second Leq ⅓-octave sound levels
  - Wind speed and temperature measurements every 10 seconds
  - Digital audio recordings (64 kbps MP3)

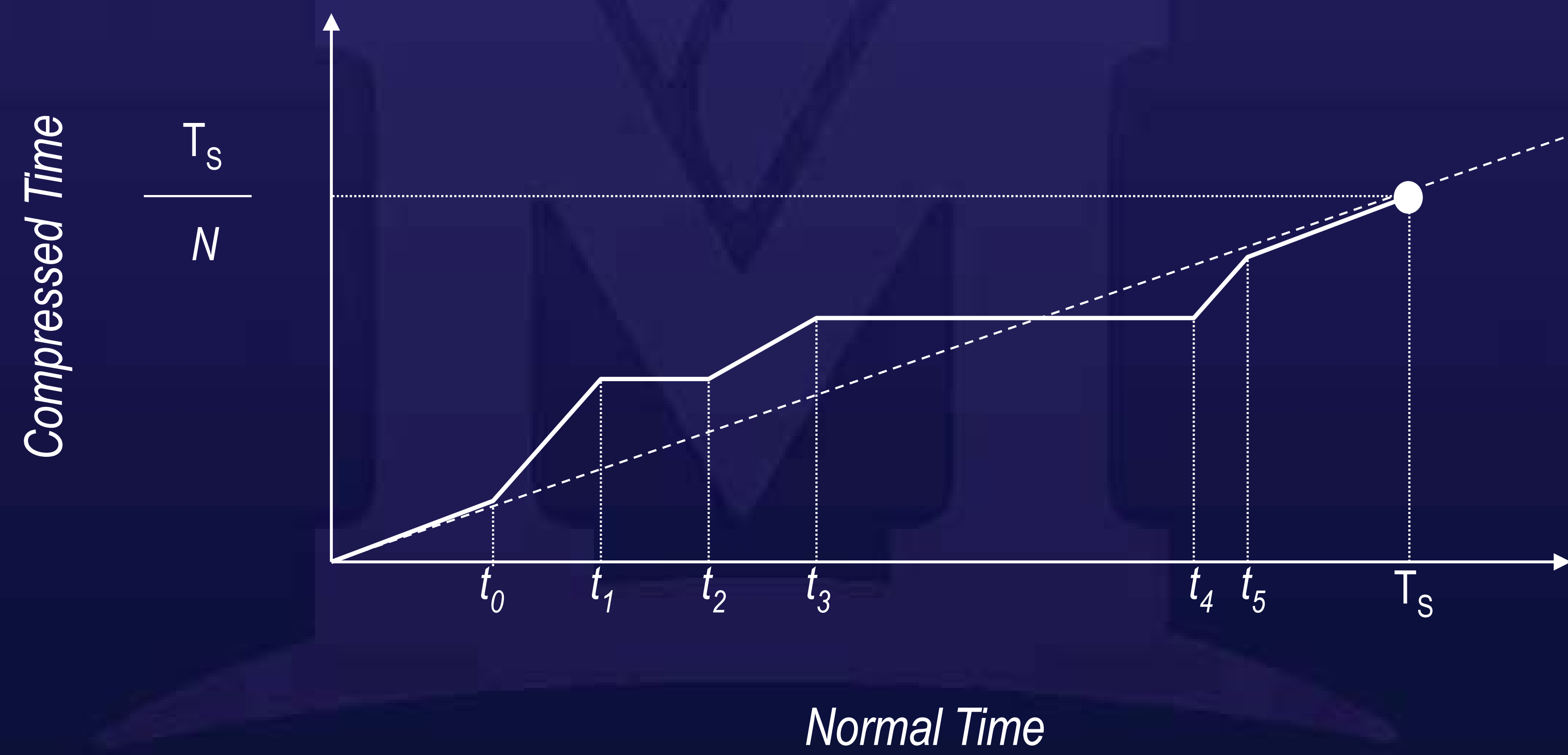
- 8,760 hours of audio


1/3 Octave Spectrogram for GRKO on 2009-05-04 (Unweighted)
Frequency (Hz)
4000
400
40
00
01
02
03
04
05
06
07
08
09
10
11
12
13
14
15
16
17
18
19
20
21
22
23
:05 :10 :15 :20 :25 :30 :35 :40 :45 :50 :55
Time (min)
-9 -6 -3 0 3 6 9 12 15 18 21 24 27 30 33 36 39 42 45 48 51 54 57 60 63 66 69 72 75 78 81 84 87 90
Sound Pressure Level (dB)

Why extreme time scale compression?

- Length of recordings is unwieldy: automated playback assistance is needed
- Researchers and the public would like a meaningful sampling
- Simple block-downsampling may not capture sonic *texture* effectively

Non-uniform time warp concept



Options

- Formal time-scale compression
- Manual editing
- Block downsampling
- Transition and event-based selection
- Computational Auditory Scene Analysis?

Texture Retention?

Example: compress 24 hours of audio into 12 minutes (N=120)

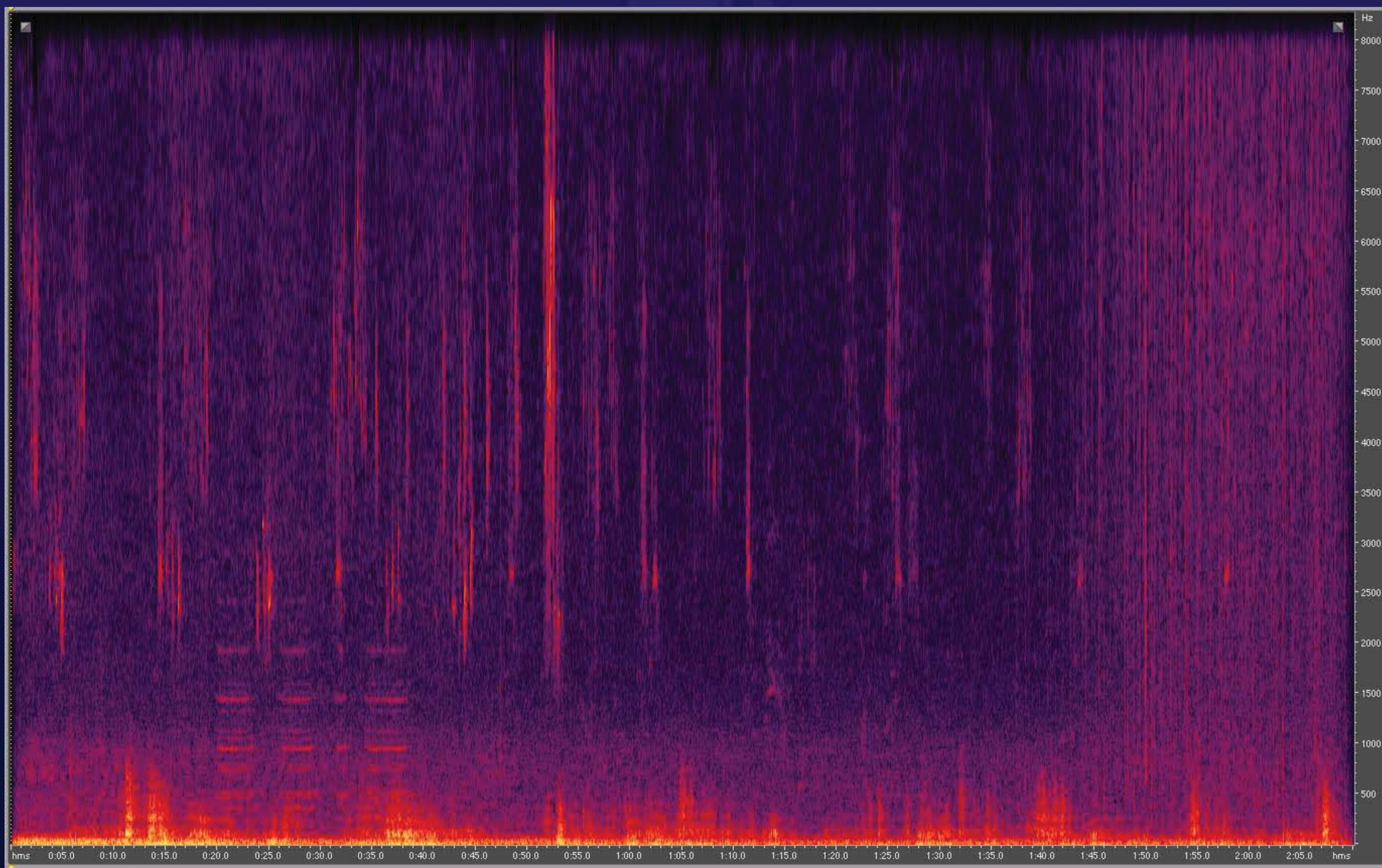
- Choose 1 second of sound every two minutes?
- Choose the most representative 1 second of sound from every two minute interval?
- Strategy: we have 720 seconds to “cover” 86,400 seconds, so “choose wisely.”

Sonic Event Map

- Identify *transitions* in the background sound
- Locate intermittent foreground sounds
- Allocate available coverage in the “optimum” manner within the constraints

Example Spectrogram

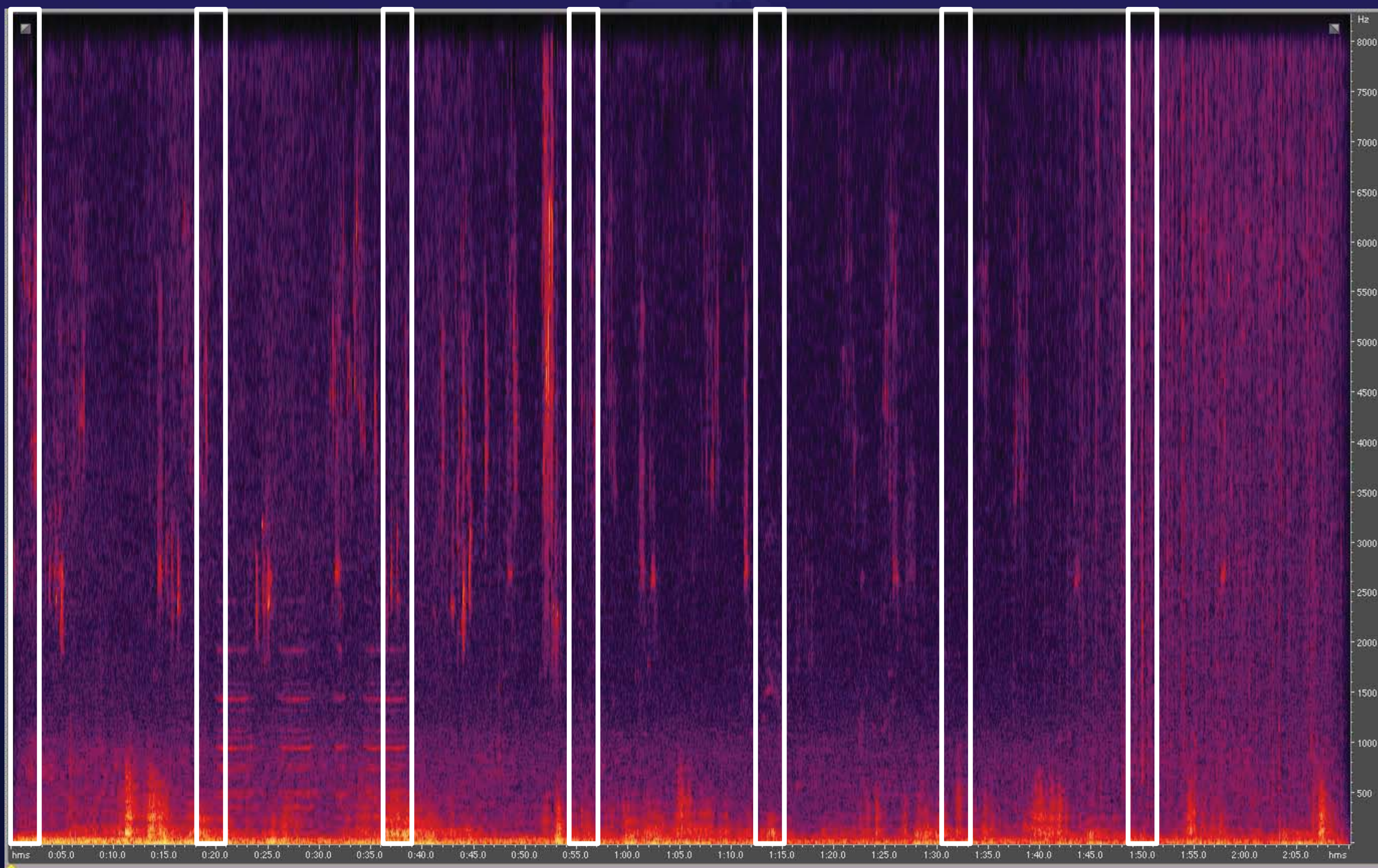
8kHz



0

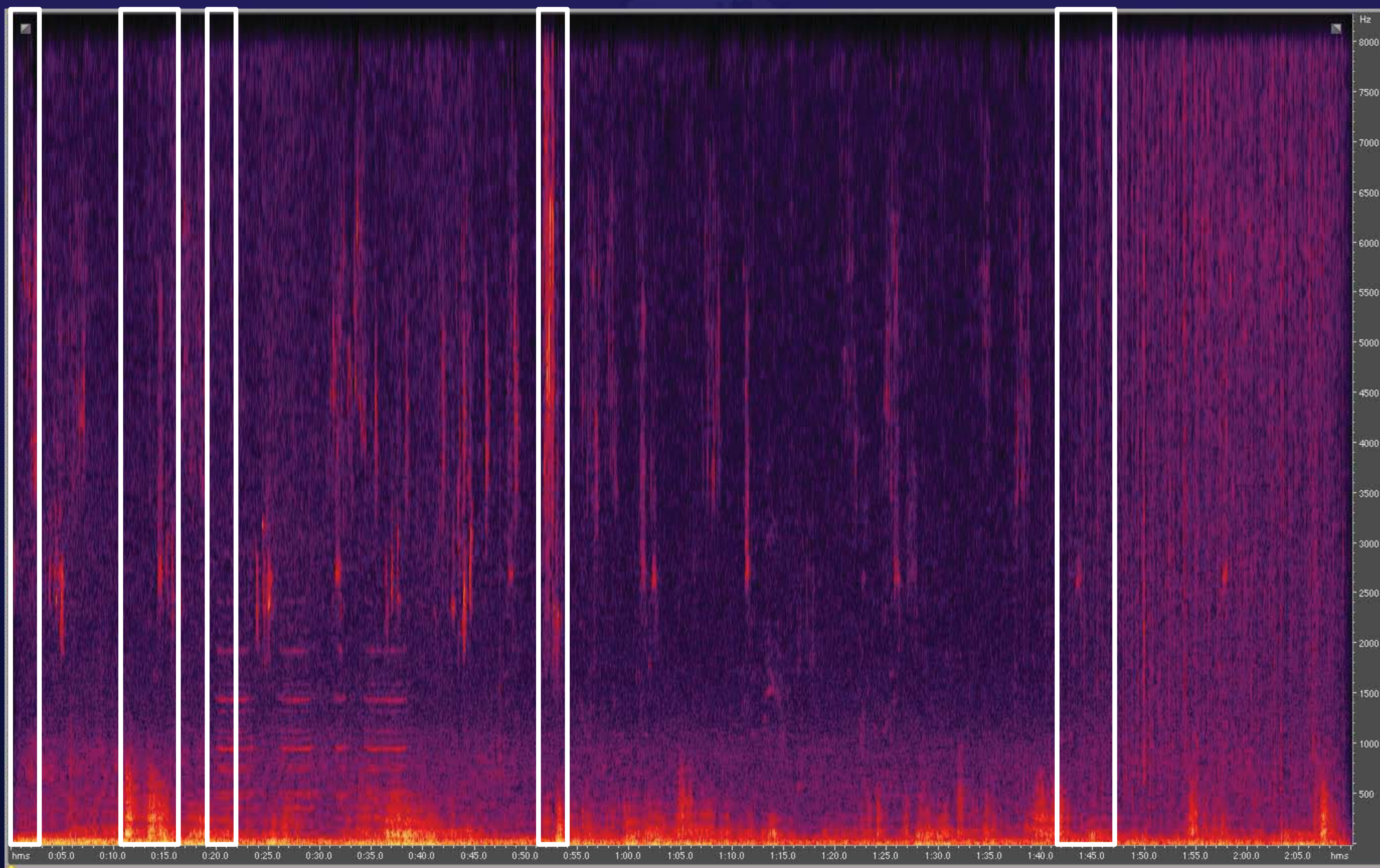
2’ 10”

Uniform N=7 ?



Example Transition Map (N=7)

8kHz

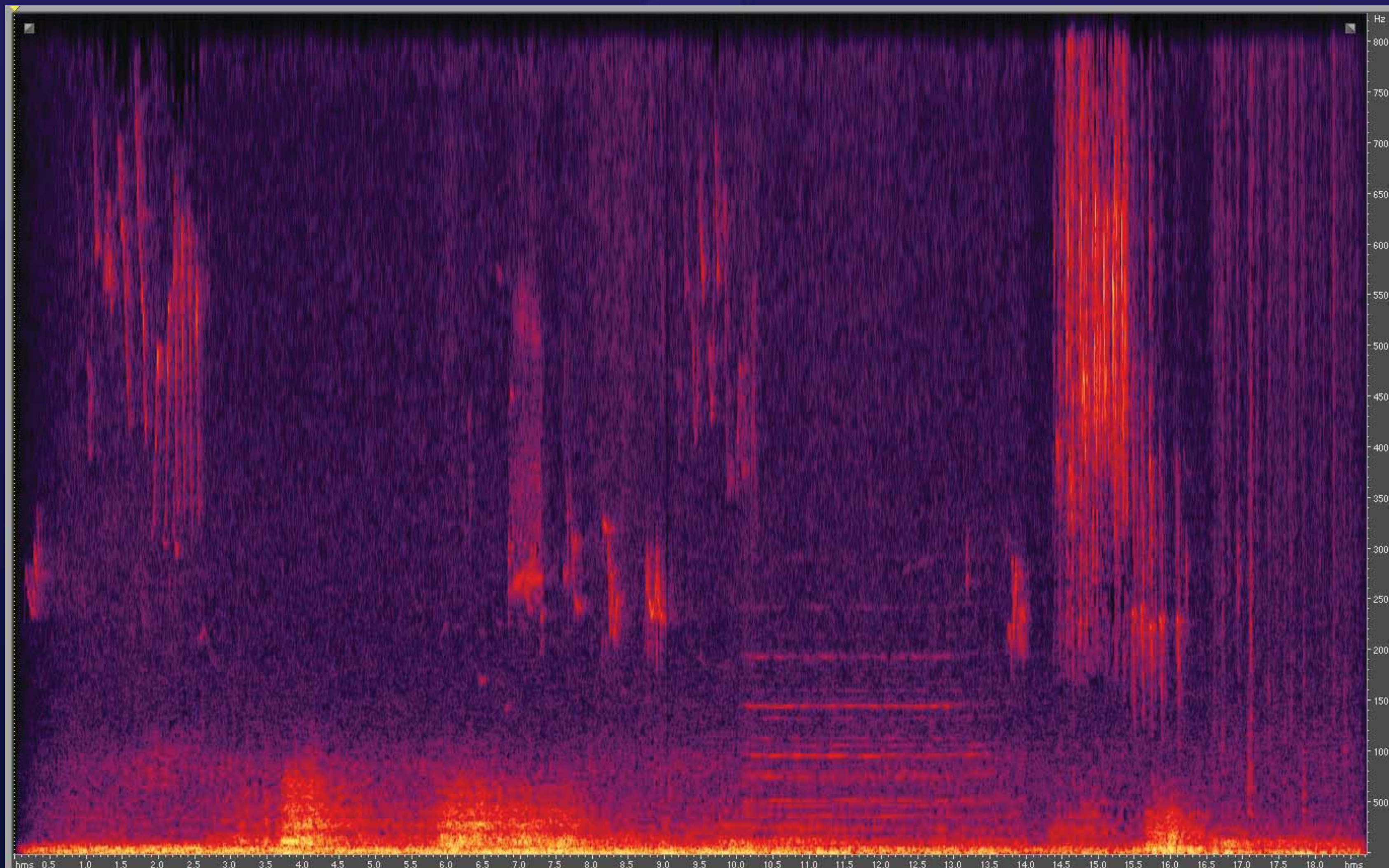


0

2' 10"

Reconstructed Signal (N=7)

8kHz



0

18”

Conclusions and Future Work

- General strategy works well, but reliability is not yet guaranteed
- Desire automatic segmentation and sound source identification
- Need an objective formula for texture determination and classification
- Key realization: *this is really a subjective data compression problem*

Thank you for your attention.

http://ece.montana.edu/rmaher/audio_monitor/grko.htm

Acknowledgements

- **Grant-Kohrs Ranch NHS, Deer Lodge, MT**
  - Christine Ford, Integrated Resources Program Manager
- **NPS Natural Sounds Program, Ft. Collins, CO**
  - Kurt Fristrup, Emma Lynch, Damon Joyce
- **Rocky Mountain Cooperative Ecosystem Studies Unit (RM-CESU), Missoula, MT**
  - Kathy Tonnessen, Research Coordinator
  - Lisa Gerloff, Executive Coordinator

Cultural soundscape of the Grant-Kohrs Ranch NHS

Robert C. Maher

Electrical and Computer Engineering

Montana State University - Bozeman

MONTANA STATE UNIVERSITY | College of ENGINEERING

Mountains & Minds

Outline

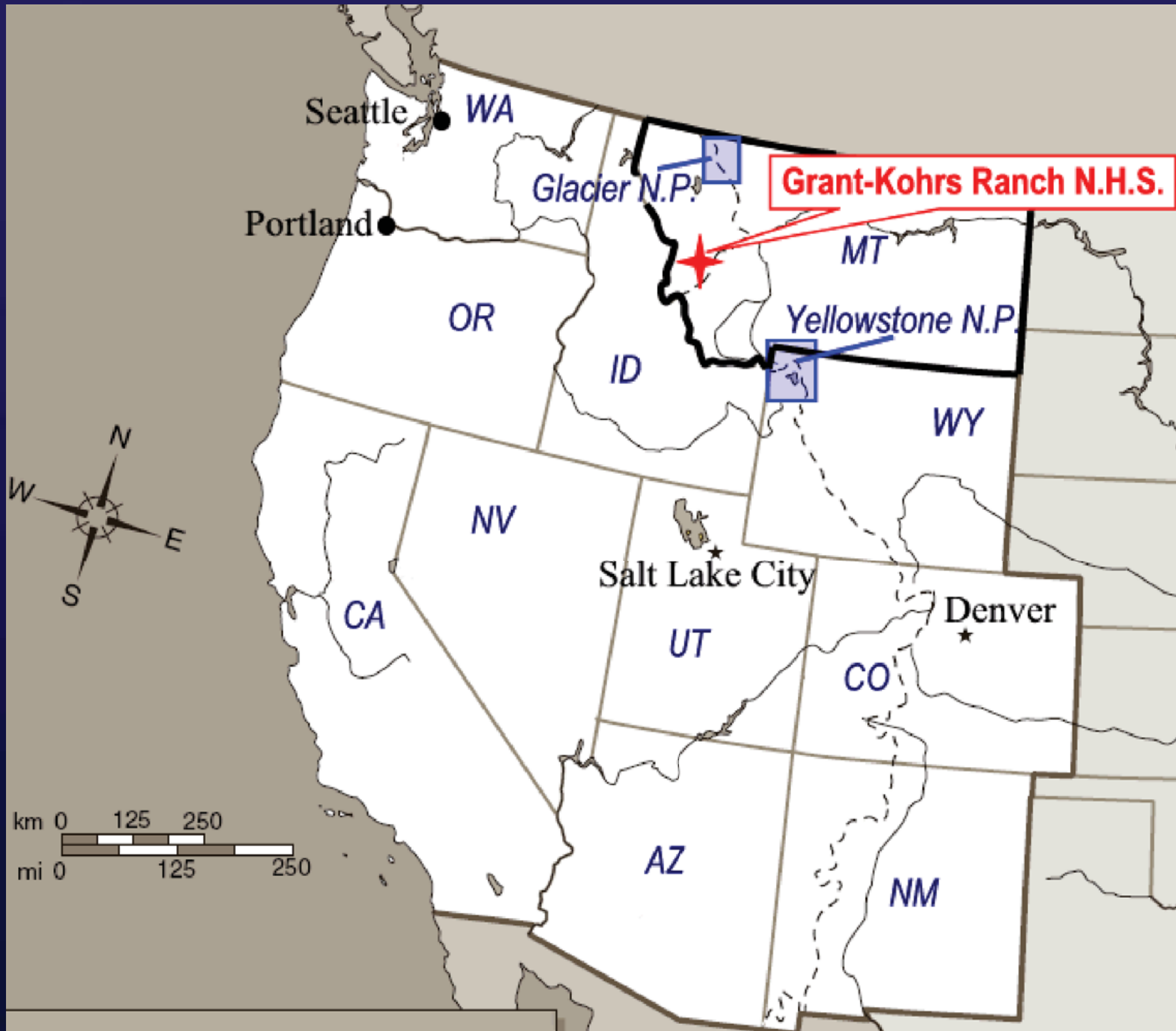
- Introduction
- Long-term acoustical acquisition
  - Sound level meter
  - Audio recorder
  - Anemometer
- Assessment and analysis
  - Biophony, Geophony, Anthrophony
  - 1/3rd octave band analysis
  - MP3 recording 8,760 hours long: automated analysis
- Conclusions

Introduction

- Grant-Kohrs Ranch (GRKO) history
- NPS sound monitor history
- GRKO soundscape: challenges and needs
- Acoustical Plan

Grant-Kohrs Ranch National Historic Site (1977)

- Deer Lodge, Montana
- A working cattle ranch commemorating the heritage of American cowboys, stock growers, and cattle operations during the 19th and 20th centuries.
- Congress: maintain the site as a working ranch.
- Cultural soundscape is essential: all the sights, sounds, and sensations associated with ranching.


Grant-Kohrs Ranch N.H.S.
Glacier N.P.
Yellowstone N.P.
Seattle
Portland
Salt Lake City
Denver
WA
OR
ID
MT
WY
NV
CA
UT
CO
AZ
NM
N
E
S
W
km 0 125 250
mi 0 125 250




MONTANA STATE UNIVERSITY | College of ENGINEERING

Mountains & Minds

Soundscape

Three Sonic Components (*Krause*):

- ***Biophony*** -- animal and biological sounds
- ***Geophony*** -- geological, hydrological, and meteorological sounds
- ***Anthrophony*** -- sounds caused by humans and human activity

Grant-Kohrs Challenges

Substantial increase in transportation noise may impact the integrity of the ranch's cultural soundscape.

- Highway traffic noise: I-90 passes within 1 km of the GRKO Visitor Center
- Deer Lodge airport general aviation expansion (2.5 km southwest of the GRKO Visitor Center)
- Neighboring ranch was purchased, subdivided into luxury home sites. Some homeowners may now fly into the area in private jets and helicopters.
- Potential establishment of a rifle range in the vicinity of the ranch.

Site elevation: 1370 meters (~4500 ft) ASL
Coordinates:
46°24'27.24"N 112°44'27.60"W

National Park Service Act (1916)

- *"...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."*

Soundscape Regulatory Context

- 1872 Yellowstone National Park Act
- 1916 National Park Service (NPS) Organic Act
- 1949 Executive Order 10092 (Boundary Waters no-fly zone)
- 1964 Wilderness Act
- 1969 National Environmental Policy Act
- 1972 Noise Control Act
- 1987 National Parks Overflights Act (NPOA)
- 1988 Special Federal Aviation Regulation (SFAR) 50-2 (GRCA)
- 2000 National Parks Air Tour Management Act
- 2000 NPS Director's Order #47 (soundscape preservation)
- 2002 Winter Use Plan (Yellowstone)
- 2006 NPS Management Policies (soundscapes)

- Miller, Nicholas, P., "US National Parks and management of park soundscapes: a review,“ Applied Acoustics, vol. 69(2), pp. 77-92, February 2008
- R.C. Maher, J. Gregoire, and Z. Chen, "Acoustical monitoring research for national parks and wilderness areas," Preprint 6609, Proc. 119th Audio Engineering Society Convention, New York, NY, October 2005.

NPS Management Policies 2006

National Park Service *Management Policies 2006* include natural and cultural sound resources within park units.

- 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

Engineering Considerations

- Long-term soundscape monitoring and statistical assessment (24/7/365)
- Low power audio recording equipment suitable for harsh environments
- Cost appropriate for widespread use
- Calibration and stability for both ecological research and regulatory monitoring

Acoustical Plan

- GRKO does not currently have any soundscape data
- A “baseline” is needed to document the seasonal and diurnal soundscape to enable management under 2006 NPS policies
- Automated measurement of sound levels and sound recordings continuously for 365 days

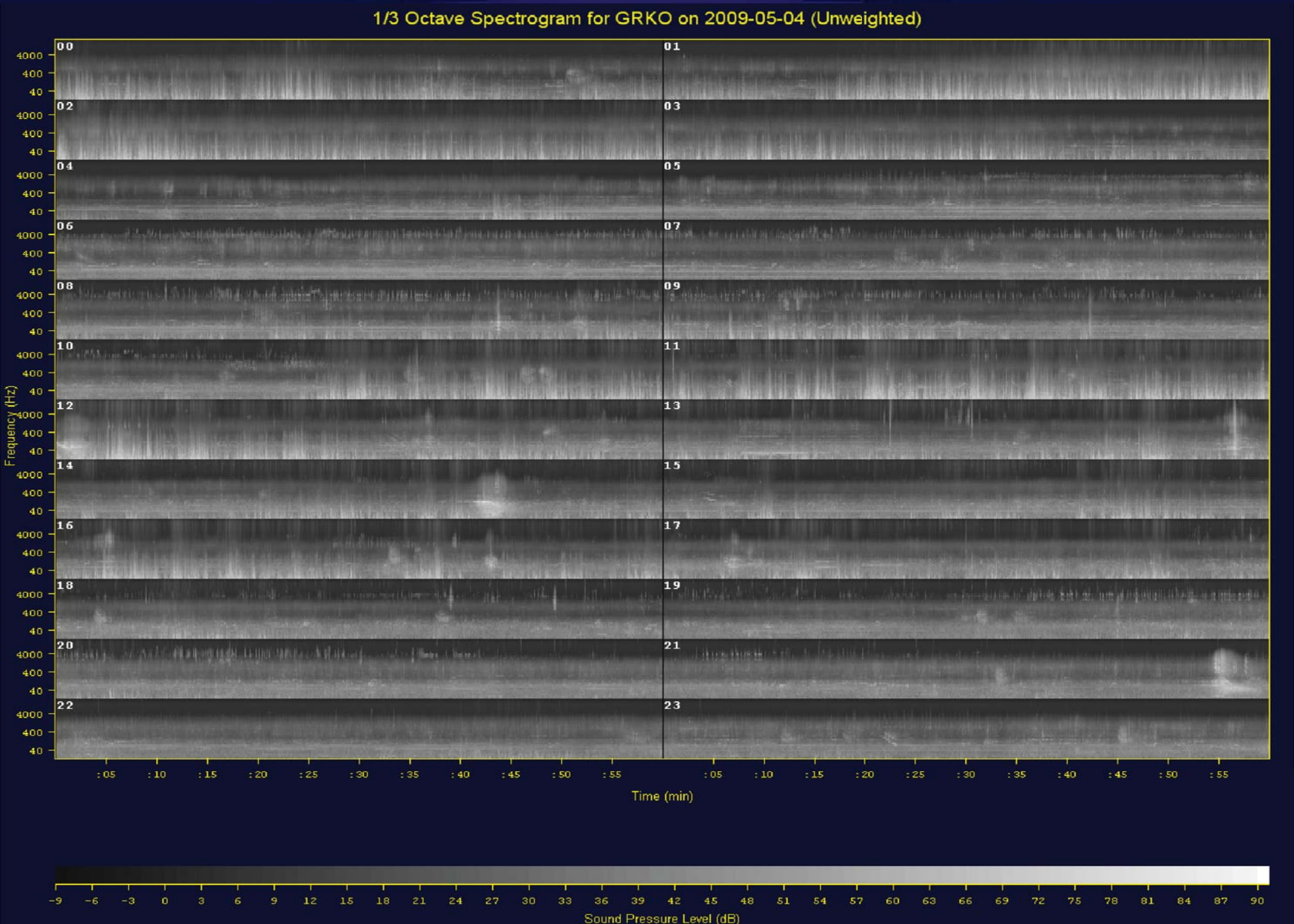
NPS Technical Requirements

- Wind speed and system temperature measurements logged automatically every 10 seconds.
- 1 second Leq ANSI Type 1 ⅓-octave sound levels (~427 MB per month)
- Digital audio (MP3) recordings: continuous 64kbps from sound level meter microphone (~25 GB per month)

Long-Term Collection

Project Status

- Data collection complete: one full year!
- Primary sounds:
  - Biophony: birds and occasional livestock
  - Anthrophony: distant highway traffic and nearby freight trains
  - Geophony: primary issue is wind
    - 10-40 km/hr not uncommon
    - Turbulence and audio signal cut-outs


1/3 Octave Spectrogram for GRKO on 2009-05-04 (Unweighted)
Frequency (Hz)
4000
400
40
00 01 02 03 04 05 06 07 08 09 10 11 12 13 14 15 16 17 18 19 20 21 22 23
:05 :10 :15 :20 :25 :30 :35 :40 :45 :50 :55
Time (min)
-9 -6 -3 0 3 6 9 12 15 18 21 24 27 30 33 36 39 42 45 48 51 54 57 60 63 66 69 72 75 78 81 84 87 90
Sound Pressure Level (dB)

Research Tasks

- Automated analysis procedures
  - Correlation of SPL, MP3, and wind data
  - Fast search of MP3 spectral data
  - Assisted identification of bird songs, etc.
- Integrity and sustainability of collection procedure
- Permanent archiving for subsequent use

Conclusions

- Basic data collection framework is useful
- Automated data analysis is essential
- Prefer higher quality uncompressed audio data
- Wind issue
- Site setup and maintenance issues

Acknowledgements

- **Grant-Kohrs Ranch NHS, Deer Lodge, MT**
  - Christine Ford, Integrated Resources Program Manager
- **NPS Natural Sounds Program, Ft. Collins, CO**
  - Kurt Fristrup, Emma Lynch, Damon Joyce
- **Rocky Mountain Cooperative Ecosystem Studies Unit (RM-CESU), Missoula, MT**
  - Kathy Tonnessen, Natural Resources Research Coordinator
  - Lisa Gerloff, Executive Coordinator

Sound Examples

- March 18, 2009 9:34PM MDT (45")
- April 15, 2009 6:13AM MDT (before dawn)(1')
- May 1, 2009 11:22AM MDT (5')
- May 4, 2009 6:23AM MDT (after dawn) (2.5')
- July 6, 2009 ~noon (6.5')
- Dec 30, 2009 9:30PM MST (2.5')
- July 6, 2009 ~1:30PM MDT (2')

- http://ece.montana.edu/rmaher/audio_monitor/grko.htm

Example 1: 2009 April morning, just before dawn

Example 2: 2009 May morning, just after dawn

090504_4.17.21.276.mp3

Example 3: 2009 December evening, 9:30PM

GRKO_20091230_2135_coyotes.MP3

Example 4: 2009 July, around noon

GRKO_20090706_noon_storm.MP3

Example 5: 2009 July, mid-afternoon

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.

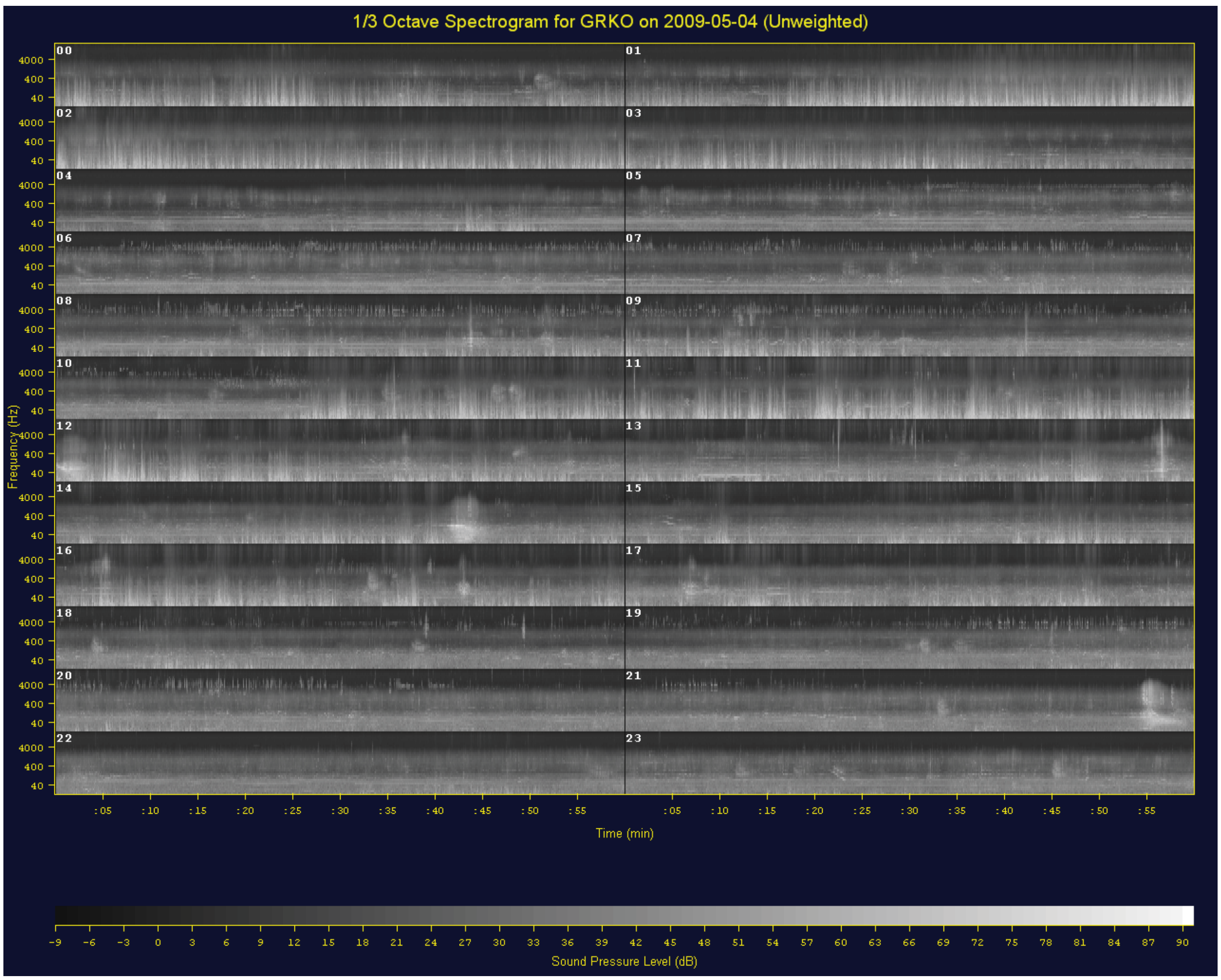


Figure 3: 24 hour spectrographic summary of the GRKO monitor site on 2009 May 4.

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.

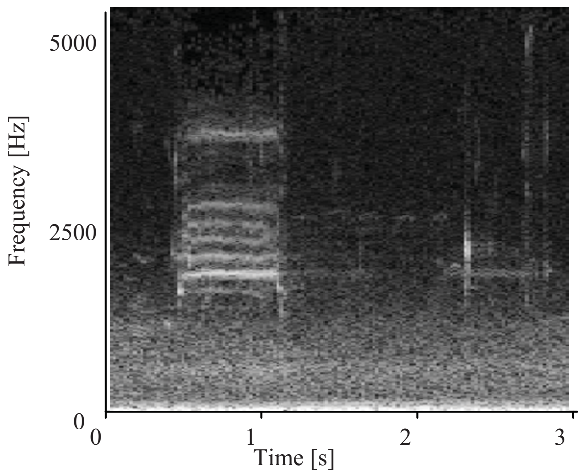


Figure 5: Bird vocalization: Red-winged blackbird (*Agelaius phoeniceus*)

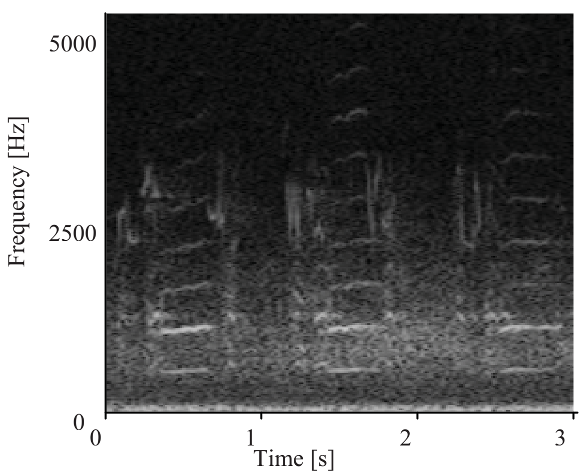


Figure 6: Canada goose (*Branta canadensis*)

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.

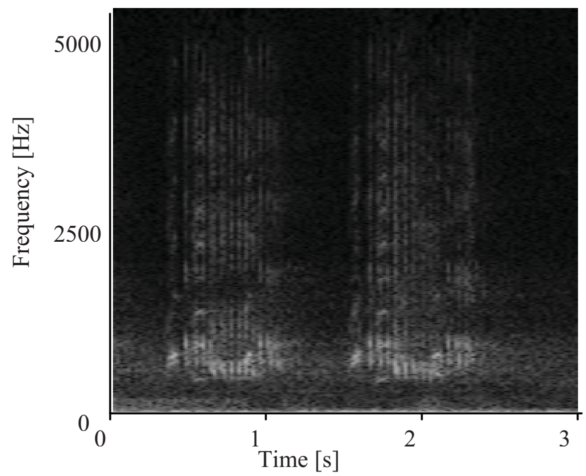


Figure 7: Sandhill crane (*Grus canadensis*)

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.

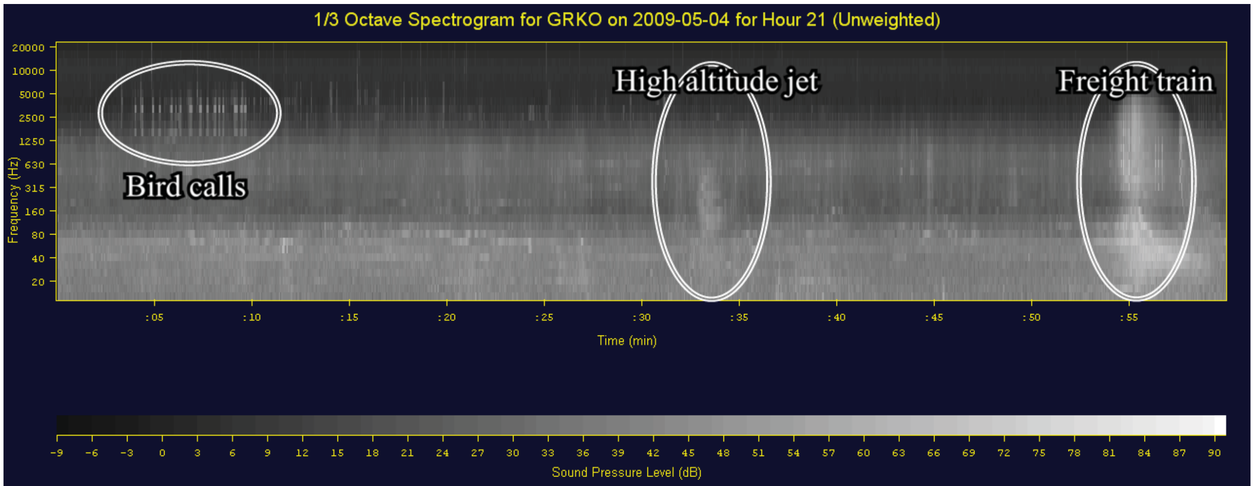


Figure 8: Example display showing the 1/3rd octave spectrogram for one hour of sound

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 $1/3^{rd}$ 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.

Baseline Sound Monitoring at Grant-Kohrs Ranch NHS

Robert C. Maher

Electrical and Computer Engineering

Montana State University - Bozeman

Outline

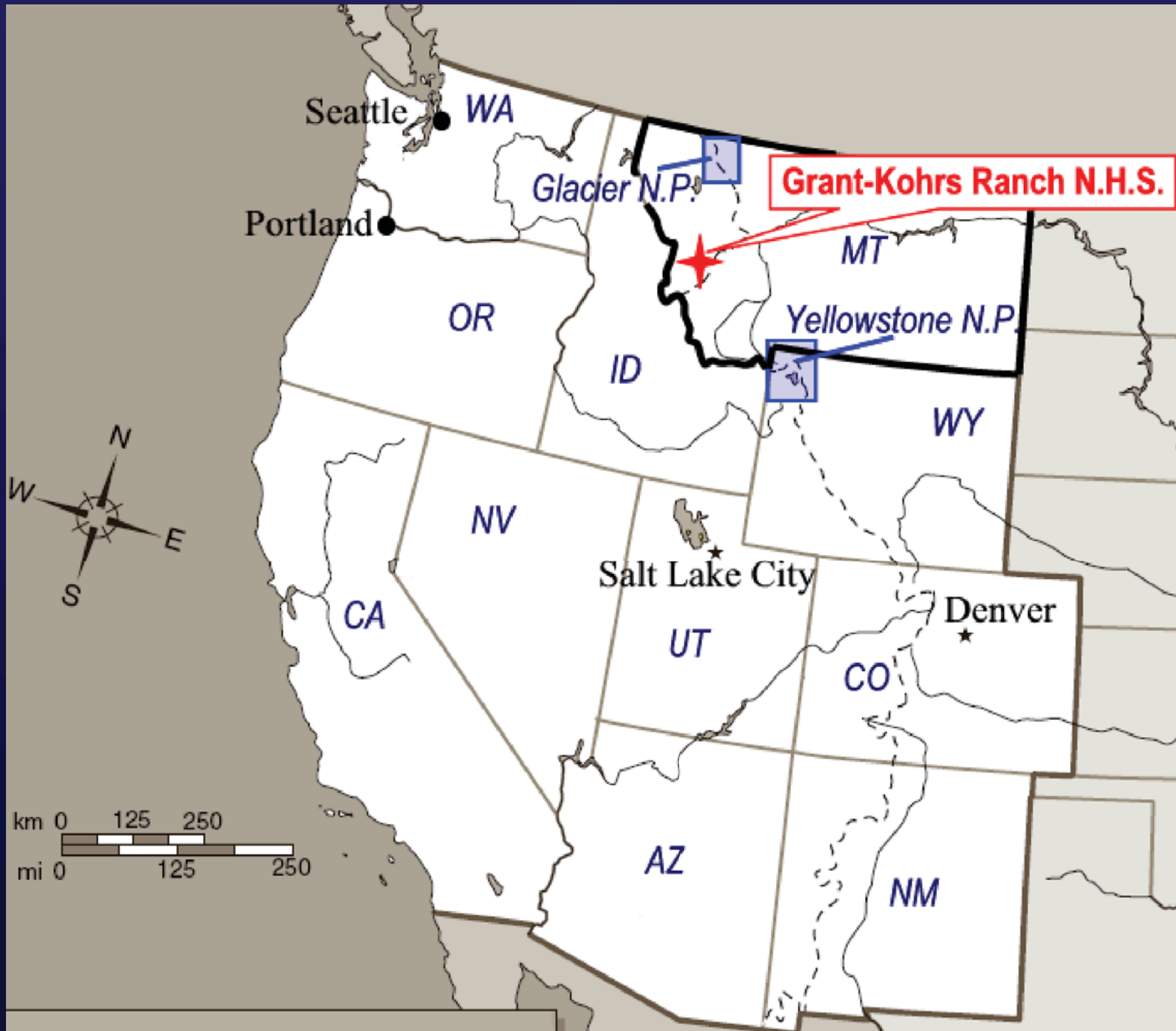
- Introduction
- Long-term acoustical acquisition
  - Sound level meter
  - Audio recorder
  - Anemometer
- Assessment and analysis
  - Biophony, Geophony, Anthrophony
  - 1/3rd octave band analysis
  - MP3 recording 8,760 hours long: automated analysis
- Conclusions

Introduction

- Grant-Kohrs Ranch (GRKO) history
- NPS sound monitor history
- GRKO challenges and needs
- Acoustical Plan

Grant-Kohrs Ranch National Historic Site (1977)

- Deer Lodge, Montana
- A working cattle ranch commemorating the heritage of American cowboys, stock growers, and cattle operations during the 19th and 20th centuries.
- Congress: maintain the site as a working ranch.
- Cultural soundscape is essential: all the sights, sounds, and sensations associated with ranching.


Grant-Kohrs Ranch N.H.S.
Seattle
WA
Portland
Glacier N.P.
MT
Yellowstone N.P.
OR
ID
WY
N
W
E
S
NV
Salt Lake City
Denver
CA
UT
CO
km 0
125
250
mi 0
125
250
AZ
NM




MONTANA STATE UNIVERSITY | College of ENGINEERING

Mountains & Minds

Soundscape

Three Sonic Components (*Krause*):

- ***Biophony*** -- animal and biological sounds
- ***Geophony*** -- geological, hydrological, and meteorological sounds
- ***Anthrophony*** -- sounds caused by humans and human activity

Grant-Kohrs Challenges

Substantial increase in transportation noise may impact the integrity of the ranch's cultural soundscape.

- Highway traffic noise: I-90 passes within 1 km of the GRKO Visitor Center
- Deer Lodge airport general aviation expansion (2.5 km southwest of the GRKO Visitor Center)
- Neighboring ranch was purchased, subdivided into luxury home sites. Some homeowners may now fly into the area in private jets and helicopters.
- Potential establishment of a rifle range in the vicinity of the ranch.

Site elevation: 1370 meters (~4500 ft) ASL
Coordinates:
46°24'27.24"N 112°44'27.60"W

National Park Service Act (1916)

- *"...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."*

Soundscape Regulatory Context

- 1872 Yellowstone National Park Act
- 1916 National Park Service (NPS) Organic Act
- 1949 Executive Order 10092 (Boundary Waters no-fly zone)
- 1964 Wilderness Act
- 1969 National Environmental Policy Act
- 1972 Noise Control Act
- 1987 National Parks Overflights Act (NPOA)
- 1988 Special Federal Aviation Regulation (SFAR) 50-2 (GRCA)
- 2000 National Parks Air Tour Management Act
- 2000 NPS Director's Order #47 (soundscape preservation)
- 2002 Winter Use Plan (Yellowstone)
- 2006 NPS Management Policies (soundscapes)

- Miller, Nicholas, P., "US National Parks and management of park soundscapes: a review," Applied Acoustics, vol. 69(2), pp. 77-92, February 2008
- R.C. Maher, J. Gregoire, and Z. Chen, "Acoustical monitoring research for national parks and wilderness areas," Preprint 6609, Proc. 119th Audio Engineering Society Convention, New York, NY, October 2005.

NPS Management Policies 2006

National Park Service *Management Policies 2006* include natural and cultural sound resources within park units.

- 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

Engineering Considerations

- Long-term soundscape monitoring and statistical assessment (24/7/365)
- Low power audio recording equipment suitable for harsh environments
- Cost appropriate for widespread use
- Calibration and stability for both ecological research and regulatory monitoring
- Automatic detection and recognition of sound sources

Acoustical Plan

- GRKO does not currently have any soundscape data
- A “baseline” is needed to document the seasonal and diurnal soundscape to enable management under 2006 NPS policies
- Automated measurement of sound levels and sound recordings continuously for 365 days

NPS Technical Requirements

- Wind speed and system temperature measurements logged automatically every 10 seconds.
- 1 second Leq ANSI Type 1 ⅓-octave sound levels (~427 MB per month)
- Digital audio (MP3) recordings: continuous 64kbps from sound level meter microphone (~25 GB per month)

Long-Term Collection


March 17, 2009

May 5, 2009


Project Status

- So far, two months of data collected (mid-March to early May)
- Primary sounds:
  - Biophony: birds and occasional livestock
  - Anthrophony: distant highway traffic and nearby freight trains
  - Geophony: wind, Wind, WIND!
    - 10-40 km/hr not uncommon (Beaufort 5 or 6)
    - Turbulence and audio signal cut-outs

Research Tasks

- Typical aural sampling protocol
  - Select one 24 hour period per “interval”
  - Audition 10 sec of audio taken every 2 minutes, manually documenting any recognizable sonic events
  - 10 seconds of audio every 2 minutes comprises 720 sample segments, or two hours of extracted audio.
  - Experts needed to identify biological sounds.

And what about the remaining 22 hours that day…and all the other 24 hour periods??

Research Tasks (cont.)

- Automated analysis procedures
  - Correlation of SPL, MP3, and wind data
  - Fast search of MP3 spectral data
  - Assisted identification of bird songs, etc.
- Integrity and sustainability of collection procedure
- Permanent archiving for subsequent use

Sound Examples

- March 18, 2009 9:34PM MDT (45”)
- April 15, 2009 6:13AM MDT (before dawn)(1’)
- May 1, 2009 11:22AM MDT (5’)
- May 4, 2009 6:23AM MDT (after dawn) (2.5’)

- http://ece.montana.edu/rmaher/audio_monitor/grko.htm

Conclusions

- Basic data collection framework is useful
- Automated data analysis is essential
- Prefer higher quality uncompressed audio data
- Wind issue
- Site setup and maintenance issues

Acknowledgements

- **Grant-Kohrs Ranch NHS, Deer Lodge, MT**
  - Christine Ford, Integrated Resources Program Manager
- **NPS Natural Sounds Program, Ft. Collins, CO**
  - Kurt Fristrup, Emma Lynch, Damon Joyce
- **Rocky Mountain Cooperative Ecosystem Studies Unit (RM-CESU), Missoula, MT**
  - Kathy Tonnessen, Natural Resources Research Coordinator
  - Lisa Gerloff, Executive Coordinator

Audio Engineering Society

Convention Paper 6609

Presented at the 119th Convention
2005 October 7–10 New York, New York USA

Acoustical Monitoring Research for National Parks and Wilderness Areas

Robert C. Maher, B. Jerry Gregoire, and Zhixin Chen

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

ABSTRACT

The natural sonic environment, or *soundscape*, of parks and wilderness areas is not yet fully characterized in a scientific sense. Published research in the U.S. National Park System is generally based on short-term sound level measurements or visitor response surveys associated with regulatory evaluation of noise intrusions from motorized recreational vehicles, tour aircraft, or nearby industrial activity. This paper reviews the history of soundscape studies in the National Park System and describes several recent advances that will allow automated recording and analysis of long-term audio recordings covering days, weeks, and months at a time.

1. INTRODUCTION

When the first National Park (Yellowstone) was created by the U.S. Congress in 1872, the telephone and the phonograph had not yet been demonstrated. Powered aircraft, electrical amplifiers, and sensitive microphones were more than 30 years away. Few, if any, could conceive a time in the future that there would be scientific studies of naturally occurring sounds, the impact of human activity, and the issue of sound conservation.

From the start of the 20th century visitors to U.S. National Parks have expected to find unique natural features, sites of historical significance, wildlife living in a natural state, and lands set aside:

> "...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]

The inherent tension between the need to "provide for the enjoyment" of current visitors while leaving the Parks "unimpaired for the enjoyment of future generations" has kept park managers on their toes ever since.

Among the features identified by the National Park Service (NPS) for protection and monitoring is the acoustical environment, or *natural soundscape*, of each

This paper includes material from the unpublished white paper: R.C. Maher, "Obtaining Long-Term Soundscape Inventories in the U.S. National Park System," January 2004.

park [2]. The natural soundscape refers to the intrinsic acoustical environment of an area without the presence of human-caused sound. Similar terms include *natural quiet* and *natural sound environment*.

The three basic constituents of the soundscape in a National Park comprise the *biophony* (animal sounds), the *geophony* (geological, hydrological, or meteorological sounds), and the *anthrophony* (sounds caused by humans) [3]. It is important to understand that natural quiet does not imply silence; rather it implies that only the natural sound sources are present. For example, the subtle sound of wind blowing through a forest, the babble of water in a stream, and the vocalizations of birds, amphibians, and other animals are all understood to be features of the natural soundscape. Yet equally valid components of natural quiet may include loud sounds like the rumble of an avalanche, the howling wind and cracking thunder during a summer storm, the crash of ocean waves, and the powerful roar of a waterfall.

As a unique natural feature, the soundscape of a National Park must be evaluated in some scientific manner so that NPS may know the current baseline and establish whether or not there is a trend in the natural sonic environment. If change is occurring, the park managers need to determine if the change is due to some natural process (seasons, migrations, climate, etc.), or due to changes in park use (number of visitors, construction of new facilities, motorized recreation, etc.). Although there is general agreement that objective and scientific acoustical monitoring is desirable, there remain many questions regarding how best to locate the monitoring sites, collect the data, evaluate the results, distinguish between trends and natural sonic variations, and determine what management actions, if any, are required to provide soundscape preservation or restoration.

The remainder of this paper is divided into sections that cover the history of sound surveys in the U.S. National Park System, describe the need for future studies, and give some examples of data collection and off-line processing.

2. HISTORY: SOUND AND NOISE SURVEYS ON U.S. PUBLIC LANDS

Over the years the major studies of soundscapes in the National Parks have been posed in the context of man-made noise intrusion. Published reports and papers regarding sound in public parks and wilderness areas began to appear in the 1970s and 1980s. Prior to that time it is clear that park visitors and managers understood the notion of natural quiet and valued it as a park resource along with fresh air and clean water, but apparently little if any formal research was conducted. Nevertheless, the connection between the natural landscape and the natural soundscape was observed by many influential naturalists.

> You feel the absence of sound—the oppression of absolute silence… But as it is, the spirit of man sympathizes with the deep gloom of the scene, and the brain reels as you gaze into this profound and solemn solitude.
>
> *Excerpt from Nathaniel Pitt Langford's 1905 publication regarding the 1870 Washburn Expedition's arrival at the Grand Canyon of the Yellowstone [4, pp. 30-32].*

> But for the time being, around my place at least, the air is untroubled, and I become aware for the first time today of the immense silence in which I am lost. Not a silence so much as a great stillness—for there are a few sounds: the creak of some bird in a juniper tree, an eddy of wind which passes and fades like a sigh, the ticking of the watch on my wrist—slight noises which break the sensation of absolute silence but at the same time exaggerate my sense of the surrounding, overwhelming peace.
>
> *Edward Abbey on the sounds of Arches National Monument in the late 1950s [5, p. 11].*

> And listen again to its sounds: get far enough away so that the noise of falling tons of water does not stun the ears, and hear how much is going on underneath—a whole symphony of smaller sounds, hiss and splash and gurgle, the small talk of side channels, the whisper of blown and scattered spray gathering itself and beginning to flow again, secret and irresistible, among the wet rocks.
>
> *Wallace Stegner recalling an experience of his youth near a mountain stream [6, pp. 42-43].*

> Silence belongs to the primitive scene. Without it the vision of unchanged landscape means little more than rocks and trees and mountains. But with silence it has significance and meaning. … How swiftly it changes if all natural sounds are replaced by the explosive violence of combustion engines and speed. At times on quiet waters one does not speak aloud, but only in whispers, for at such moments all noise is sacrilege.
>
> *Sigurd F. Olson considering the Quetico-Superior canoe country [7, pp. 51-52].*

2.1. The Early Days

As early as the 1940s the issue of wilderness soundscape preservation was reflected in President Harry Truman's Executive Order 10092, "Establishing an Airspace Reservation Over Certain Areas of the Superior National Forest in Minnesota" [8]. The Executive Order restricted air travel in the Boundary Waters Canoe Area of northern Minnesota by an outright prohibition on float plane landings, and by banning flights of any kind lower than 4000 feet above mean sea level.

The Order represented the first time a ban had been placed on aircraft flight in order to preserve a wilderness area anywhere in the world. Although the Order's implicit intent was to eliminate commercial airborne outfitters within the wilderness boundaries, the explicit objective was to eliminate the aircraft noise. Subsequent federal rule-making restricted motorized watercraft, and ultimately the Boundary Waters area was officially declared a protected wilderness by the Boundary Waters Canoe Area Wilderness Act of 1978 [9].

The U.S. Congress established a new category of federal land protection and conservation with the passage of the Wilderness Act in 1964 [10]. The Act included in its definition of wilderness the characteristics of a landscape

> "where man himself is a visitor who does not remain…land retaining its primeval character and influence, without permanent improvements or human habitation,"

and, furthermore, land that

> "has outstanding opportunities for solitude or a primitive and unconfined type of recreation."

This definition of "wilderness" is clearly distinct from the National Park concept, which embodies the notion of "unimpaired" future enjoyment rather than the more restrictive prohibition on permanent improvements found in the Wilderness Act.

2.2. The Environmental Movement

The rise of the U.S. environmental activist movement in the 1960s included increasing public awareness of man-made noise and its effects on the natural soundscape, particularly in urban and suburban communities. However, there was minimal concern at that time within the National Park Service since noise intrusions in the parks were generally limited to automobile tourist traffic and intermittent overflights by commercial passenger aircraft at cruising altitude. Motorized recreation was not yet popular, nor was there yet much interest in commercial tour aircraft flights.

The U.S. Defense Department sponsored studies of acoustic overpressure effects due to above-ground nuclear blasts during the 1950s. Such tests were only conducted in remote locations of Nevada and the Pacific Islands far from national parks, but the acoustical experiments showed the potential for noise from explosions to propagate over hundreds of kilometers due to atmospheric refraction and focusing [11].

In the 1960s and early 1970s the planning process in the United States for a civilian supersonic transport (SST) program led to a variety of environmental studies. Since supersonic passenger airliners would inherently create sonic booms, a number of formal studies were launched to gauge the effects of acoustic shock waves on structures, human beings, wildlife, and domestic animals. The U.S. Air Force and Federal Aviation Agency (now the Federal Aviation Administration, or FAA) conducted several sonic boom investigations, including several months of experiments with supersonic military aircraft (F-104) 21,000 to 50,000 feet over Oklahoma City, between February 3 and August 3, 1964 (Operation Bongo) [12]. The sonic booms were produced each day at pre-specified times, and population response surveys were obtained. A total of 1,254 flights were conducted. The results from similar tests during 1961-62 in St. Louis, in 1965 in Milwaukee, Chicago, and Pittsburg, and during 1966-67 at Edwards Air Force Base were similar to the extensive Oklahoma study: the tests resulted in thousands of complaints and damage claims [13]. Ultimately, the test results indicated that a significant segment of the public would not be able to ignore the sound of sonic booms—especially from the anticipated large civilian SST aircraft and their presumably unpredictable boom arrival times.

Citizen groups became active in the late 1960s, most notably the "Citizens' League Against the Sonic Boom," founded by William Shurcliff in 1967. The FAA sponsored a variety of study groups and panels, including the Supersonic Transport (SST) Community Noise Advisory Committee (July 1970), chaired by Leo

Beranek. A general consensus was developing that civilian supersonic overflights would not be tolerable public policy [12].

An 80 ton collapse of rock attributed to the sonic boom of a military aircraft was reported by a ranger at Canyon de Chelly National Monument on August 11, 1966, and a subsequent collapse was also noted on October 4. The collapse damaged a prehistoric cliff dwelling. The rangers at the park also documented 83 sonic booms between August 11 and December 22, 1966 [14]. A rockslide caused by another sonic boom blocked a road in Mesa Verde National Park on February 21, 1968 [11]. These documented incidents of damage to physical and cultural sites within the National Parks caused great concern about the potential for ongoing adverse effects of sonic booms—particularly as the U.S. Congress considered support of a civilian SST development program.

The growing significance of environmental concerns in the United States reached the Congress in the late 1960s. Congress passed the National Environmental Policy Act of 1969, requiring environmental impact statements (EIS) for all projects considered by federal agencies [15]. Legislation authorizing the creation of the Environmental Protection Agency (EPA) was passed in 1970.

Shortly thereafter, the United States Senate voted down funding for the U.S. SST effort on March 24, 1971, based on economic and environmental concerns. The FAA subsequently banned civil aircraft sonic booms. Sonic booms from military aircraft were not then nor are they now restricted by FAA rules, nor are sonic booms caused by spacecraft such as the space shuttle during atmospheric reentry. Nevertheless, the concern of the general public regarding noise and sonic boom issues has led the military to restrict supersonic operations to areas over the oceans, at high altitudes over land (above 30,000 feet), or within specific military reservations [16]. No recent examples of sonic boom damage within the U.S. National Park System have been reported, although sporadic reports of military aircraft causing audible sonic booms continue to occur.

2.3. Federal Environmental Protection

Management authority for the federal government to regulate the use of off-road motorized vehicles and boats on public lands was given in Executive Order 11644, signed by President Richard Nixon February 8, 1972 [17]. The Order specifically prohibited off-road vehicles in protected wilderness areas, and specified that off-road "areas and trails shall be located in areas of the National Park system, Natural Areas, or National Wildlife Refuges and Game Ranges only if the respective agency head determines that off-road vehicle use in such locations will not adversely affect their natural, aesthetic, or scenic values."

Federal legislation concerning community noise appeared with the Noise Control Act of 1972, signed by President Nixon on October 28, 1972 [18]. FAA retained jurisdiction over aircraft noise levels, but EPA was given advisory authority to recommend regulations. The EPA influence in the 1970s was primarily in the area of community noise, with less relevance to National Parks and wilderness areas.

In recent years the issue of personal water craft (e.g., jet skis) has been studied at Everglades (FL) and Glen Canyon (UT) National Parks. NPS has put in place extensive restrictions on where personal water craft may be used. Similarly, the noise caused by snowmobiles in the Grand Teton, Yellowstone, and the connecting roads and trails in the John D. Rockefeller, Jr., Memorial Parkway has been studied and the results have been reflected in recent regulations on allowable snowmobile sound emissions [19].

The noise of industrial facilities and mining operations near to park lands has been the subject of study and some controversy [20]. NPS has also instituted soundscape protection plans for its own maintenance and construction activities and those of the contract concessionaires on park lands to reduce, in effect, the "self noise" of the park facilities.

2.4. The Air Tour Issue

Public Law 93-620, "Grand Canyon National Park Enlargement Act", passed by the U.S. Congress in January, 1975, included language specifying the authority of the Secretary of the Interior [21]:

> "Whenever the Secretary has reason to believe that any aircraft or helicopter activity or operation may be occurring or about to occur within the Grand Canyon National Park…, including the airspace below the rims of the canyon, which is likely to cause an injury to the health, welfare, or safety of visitors to the park or to cause a significant adverse effect on the natural quiet

> and experience of the park, the Secretary shall submit to the Federal Aviation Administration, the Environmental Protection Agency pursuant to the Noise Control Act of 1972 (42 U.S.C. 4901 et seq.), or any other responsible agency or agencies such complaints, information, or recommendations for rules and regulations or other actions as he believes appropriate to protect the public health, welfare, and safety or the national environment within the park. After reviewing the submission of the Secretary, the responsible agency shall consider the matter, and after consultation with the Secretary, shall take appropriate action to protect the park and visitors."

The overlapping jurisdiction of the EPA, NPS, FAA (Department of Transportation) and other federal agencies such as the Bureau of Land Management (Department of the Interior) and the U.S. Forest Service (Department of Agriculture) created a bureaucratic tug-of-war. Congress unequivocally authorized the FAA to promulgate regulations involving all aspects of civilian aircraft operations, but conflicts between aircraft noise and community noise standards or conflicts with the requirements of the Organic Act that NPS manage the National Parks "in such manner and by such means as will leave them unimpaired for the enjoyment of future generations" were not anticipated in legislation before the 1970s.

The passage of the Grand Canyon National Park Enlargement Act in 1975 led to several preliminary studies of the noise caused by air tour overflights at the canyon [22]. The studies included calibrated sound level measurements at several locations and a variety of user surveys. This research work continued off and on for the next decade, resulting in a significant concern by NPS that the visitor experience was being degraded by aircraft noise and that the likelihood of visitor injuries or damage to the park was reaching an unacceptable level. A later study of the effects of helicopter noise on an archeological site at Grand Canyon added to this concern [23].

By 1987 the number of commercial air tour operators at the Grand Canyon grew to 40 and the annual total of individual air tour flights reached 50,000 [24, pp. 41-42].

The increasing air tour concern by NPS was tragically confirmed when a Bell Jet Ranger tour helicopter operated by Helitech, Inc., and a DeHavilland Twin Otter fixed-wing tour aircraft from Grand Canyon Airlines, Inc., collided in mid-air over the park in June, 1986, resulting in the deaths of all 25 people on board the two aircraft [25]. The accident and the existing concerns about the impact of overflight noise on Grand Canyon National Park and other units of the National Park System led the U.S. Congress to take action. The resulting legislation, Public Law 100-91, the "National Parks Overflights Act of 1987," required NPS and the U.S. Forest Service to perform studies addressing specific questions about aircraft overflight issues, and then to prepare a formal report to Congress [26]. Significantly, the legislation included funding to conduct the required surveys and research studies. Other provisions included explicit instructions on the expected restoration of natural quiet at Grand Canyon:

> "Flight-free zones are to be large areas where visitors can experience the park essentially free from aircraft sound intrusions, and where the sound from aircraft traveling adjacent to the flight-flee zone is not detectable from most locations within the zone."

The National Parks Overflights Act (NPOA) assessment studies occurred over a period of nearly eight years, culminating in the mandated report to Congress dated July, 1995. NPS hired several acoustical consulting firms to do the assessments (HMMH—Harris, Miller, Miller and Hanson, BBN—Bolt, Beranek and Newman, and Wyle Laboratories). The consultants and the Park Service managers quickly discovered that the aircraft noise standards and measurement procedures suitable for community noise studies were ill-suited to the extremely low ambient noise levels found in many national parks. For example, the sound of distant, high-altitude aircraft that would go unnoticed in an urban or suburban setting were easily audible compared to the natural quiet of the park. Similarly, noise standards based on interference with work or sleep inside a home or office did not apply very well to backpackers out on a hiking trail or sleeping in a tent. A great deal of effort was required to determine not only *what* to measure, but *how*, *where*, and for *how long* it should be measured.

Several of the major studies concluded in support of the NPOA report included experimental measurement of sound levels on the ground due to aircraft at various altitudes [27, 28, 29], dose-response surveys and questionnaires attempting to judge visitor perception and tolerance of overflight noise [30], literature reviews concerning reports on the effects of overflights on animals and cultural resources [31, 32], and surveys of air tour passengers, park service managers, and other park personnel to assess their awareness and expectations regarding overflights [33, 34].

The National Parks Overflights Act also directed the Secretary of the Interior to develop recommendations for changes in aviation regulations that would eliminate adverse effects of overflights on Grand Canyon. The Act further required the FAA to apply the Secretary's recommendations without change except in any case where the recommendation would diminish aviation safety. The Department of the Interior submitted recommendations to the FAA in December, 1987. In response, FAA issued Special Federal Aviation Regulation 50-2 [35], establishing designated flight routes, minimum altitudes for air tours, and special no-fly zones. Studies of the effectiveness of SFAR 50-2 in mitigating aircraft noise were conducted as part of the NPOA report [36]. Additional FAA rules and regulations were put in place in the late 1990s. Recent years have seen litigation by the U.S. Air Tour Association (USATA) challenging several of the flight restrictions, and also court challenges by the Grand Canyon Trust and other environmental groups urging increased flight restrictions.

2.5. The Current Situation

The U.S. Congress acted again in April, 2000, to formalize procedures for air tour operations by instituting the “National Parks Air Tour Management Act of 2000” [37]. The Act continued to split responsibility between NPS and FAA, directing that each unit of the National Park System affected by air tour flights must develop an air tour management plan. The Act also required NPS and FAA jointly to establish an advisory panel, now known as the National Parks Overflights Advisory Group (NPOAG), consisting of representatives from general aviation, commercial air tour companies, environmentalists, and Indian tribes, to provide input on the use and management of airspace over National Parks.

Various NPS units are conducting research and data collection for aircraft, personal watercraft, snowmobiles, off-highway vehicles, and other types of motorized recreation. Active data collection projects are recently concluded or currently underway at Grand Teton (WY), Glen Canyon (UT), Grand Canyon (AZ), Yellowstone (WY, ID, MT), Haleakala (HI), Denali (AK), and other park units.

3. FUTURE STUDY REQUIREMENTS

The attribute largely missing from prior studies is **a truly long-term evaluation of the natural soundscape covering all hours of the day and all seasons of the year**. Very little is known in a scientific sense about the diurnal and seasonal variations in natural sound, nor about the long-term trends in the natural soundscape. In addition to assessing human-caused sound and noise, the availability of long-term sonic data can provide a different viewpoint for studying biology and ecology within the parks.

3.1. Are Long-Term Recordings Really Needed?

For the obvious practical and cost savings, it is reasonable to consider whether occasional acoustical monitoring at a few selected locations for short periods is sufficient for most soundscape studies. For example, one might arrange to observe the soundscape at a particular location for 12 hours beginning at midnight on the first Monday of a month, and find a few dozen identifiable events such as a particular bird song, a tree snapping and falling over in the wind, a bumble bee flying by, a helicopter overflight, and the howl of a coyote. Although this information may be useful, the measurement has not necessarily provided a statistically meaningful sample of the soundscape since it is not know if the particular day is truly representative. Changes are expected due to temperature, wind, and other meteorological details, as are changes due to migration of wildlife, presence of wandering predators, growth of seasonal vegetation, etc. For some purposes, such as evaluating the characteristic sound level of a specific aircraft flight corridor at Grand Canyon, the data can be temporally sparse and still provide useful insights. On the other hand, for research involving animal population studies, correlations between sound and meteorological conditions, and discovering diurnal and seasonal trends, data must be obtained and evaluated for weeks and months at a time.

Obtaining a long-term recording of the soundscape in a particular location makes it possible to make strong and statistically significant statements about trends, wildlife observations, the impact of management decisions, and so forth.

3.2. Soundscape Instrumentation

Long-term scientific soundscape data has been difficult to obtain due to the lack of affordable, effective, and easily deployed sound monitoring gear. The equipment used for most of the existing formal sound studies is expensive precision acoustical instrumentation costing $10,000 or more per setup [38]. The gear requires special training to use. It is not intended for widespread deployment, nor is it typically designed for obtaining continuous audio signal recordings.

The NPS and Federal Aviation Administration overflight and general soundscape studies currently use a monitoring system based on a laptop computer attached to a calibrated sound level meter [39]. The 1-second L_{eq} (the mean-square sound pressure, expressed as a level in dB) and 1/3 octave sound levels calculated by the meter are collected by the laptop from a serial connection. Meanwhile, the analog signal from the sound level meter's microphone is sent to the laptop's audio input where it is sampled at 44.1 kHz with 16-bit resolution. The software continuously samples the audio data into a 20-second circular buffer. The software provides a user-selectable threshold that causes a 20-second segment of the sampled data to be saved; otherwise the buffer contents are written over. The laptop software also automatically records a 10-second sample every 4 minutes whether or not the threshold has been exceeded.

The laptop systems have been deployed for several weeks at a time and have been found to be quite reliable. However, the platform is bulky, delicate, and power-hungry. The sound level meter provides a calibrated and standardized acoustical reference, but it is expensive and perhaps redundant given the computing resources of the laptop. Power for the laptop must come from a conventional power line, or in the case of a system deployed in a remote area a photovoltaic panel and rechargeable battery must be provided [40, 41].

Furthermore, because it is difficult to anticipate all the information that may be desirable to glean from long-term sound inventories, attempting to specify acoustical preprocessing and data reduction properties in advance may be counterproductive. For example, many published acoustical studies in the National Parks have used *A-weighted* sound level measurements and one-third octave-band analyses that are considered suitable for assessing the audibility or annoyance caused by intrusive noise. The A-weighting is typically considered appropriate for noise assessment because acousticians often use weighted sound levels for comparisons and regulatory purposes. Unfortunately, the common sound level data provide insufficient information about specific natural sound sources and their temporal distribution. This shortcoming is due to the assumption of human audibility and community noise standards with the explicit goal of assessing noise intrusion rather than the general soundscape evaluation, inventory, and sound source classification.

A-weighted measurements use a frequency response filter that is intended to model the non-uniform sensitivity of the human hearing system for low-level sounds. The filter results in a degraded signal-to-noise ratio because the signal level is deliberately reduced at low and high frequencies.

Future long-term soundscape studies in the U.S. National Park System will require specialized—yet economical—test equipment. A truly useful acoustical inventory will require simultaneous monitors deployed in numerous locations for weeks and months at a time, ideally with essentially no day-to-day hands-on maintenance. Furthermore, it seems likely that the acoustical monitoring gear will be handled and deployed by non-technical personnel such as volunteers, so compact packaging, modular assembly, rugged design, and attention to human factors is vitally important.

The specialized nature of the proposed test equipment makes it difficult to count on commercial development. Rather, initially it is likely that it will be necessary to rely on custom development of prototype systems [42]. Nevertheless, the increasing availability of low-cost and low-power microelectronic components used in commercial portable products will provide an opportunity to develop custom acoustical monitors using commercially available parts.

4. PROSPECTS FOR DATA COLLECTION AND PROCESSING

Most research to date in the National Park System has focused on human audibility and annoyance of intrusive noise sources. The use of average overall and 1/3-octave levels has been found adequate for this purpose. Future research will ultimately yield automated means to identify sound sources [47, 48]. This will only be possible if continuous audio recordings are available.

The most significant research challenge will be in the interpretation and documentation of the acoustical data. Although human listeners are quite adept at detecting and classifying sounds in audio recordings, it is not sensible to assume that the hours of recordings from each monitoring system would be analyzed solely by human ears: new means for automated acoustical processing must be developed, tested, and refined. Reliable parsing of complicated audio recordings is a difficult, and remains an active research area in the field of signal processing.

These practical issues point toward two parallel research initiatives: (a) methods for acoustical data processing, and (b) prototype equipment design and evaluation.

4.1. Acoustical Data Processing

Once a multi-day audio recording has been made, the data can be returned to the research lab for analysis. To begin with, a standard set of basic measurement algorithms will be performed to characterize the sonic environment. These simple measurements will include short-time and long-time average sound pressure level, percentage of time specific levels are exceeded, and estimates of the time-variant spectral envelope of the soundscape. Moreover, since the actual audio data has been recorded, it is entirely possible to perform many different analyses on the data at any time in the future.

Next, it will be desirable to identify and classify sounds within the recording. The extreme length of the measured recordings makes analysis by human listeners essentially impossible. An automated and reliable means to detect sonic events in the hours and hours of recorded data must be invented, implemented and validated.

Identifying sound events may seem trivial since everyday experience involves many situations in which one must recognize the phone ringing, a dog barking, or rain falling on the roof, but no reliable automatic algorithms for parsing multiple concurrent sounds in an audio recording have been demonstrated. Automated detection of sound level, changes in the background noise level, and similar general features can be quite effective, but this sort of segmentation may still require considerable manual intervention. Nevertheless, the fact that the entire audio recording has been obtained allows the ongoing advances in automated audio source analysis to be used as they become available.

Significant signal processing research will be conducted to identify and classify the natural sound sources, such as animal vocalizations, flowing water, and wind interacting with vegetation and terrain. Additional research will focus on reliable methods for extracting specific sound sources of interest such as aircraft, vehicles, and other mechanical sounds. The identification and classification framework will be carried out in software using a time-frequency decomposition of the input signal followed by a stimulus matching procedure [e.g., 43, 44, 45].

4.2. Prototype Monitoring Equipment

For the prototype design/evaluation phase a rugged and self-contained monitoring platform will need to be designed and constructed. There does not appear to be a standard catalog item that satisfies the continuous long-term recording requirements, but at least one specialized product is under development by Sanchez Industrial Design, Inc. The model PADR-100 Portable Audio Data Recorder [46] is a development platform designed for long-term continuous recording (up to 7 days) with a variety of communication ports and optional accessories.

If time and cost constraints allow, it may also be feasible to develop a custom recording platform. The proposed platform would contain a digital signal processor (DSP), a calibrated microphone and data acquisition subsystem, a memory subsystem, and a power supply. The platform is intended to be deployed unattended in a remote location for at least 14 days at a time while continuously making a digital recording of the acoustical environment. Every 14 days the monitoring platform would be serviced: the recorded audio data on the hard disk drive would be brought back to the lab for subsequent off-line analysis, and the system battery would be swapped with a fresh power source. In some situations it might be possible to consider a wired or wireless data connection from the recording platform to a host computer, and perhaps rely upon remote line power so that the battery is used only for power backup purposes, but in general the platform must be designed to operate with full-time battery power and data storage for the entire 14 day period.

4.2.1. System Features

The proposed features of the prototype system include:

- Continuous wideband (at least 20Hz – 20kHz) 24-bit audio recording capability.
- Omnidirectional calibrated microphone system, ±0.5dB 20Hz – 20kHz, suitable for extremely low ambient sound levels [38].
- Overall system design capable of IEC Type 1 performance (IEC 61672-1:2002)
- Hard disk storage with at least 14 days@24 hours/day capacity.
- Low-power electronics designed for 14 day operation on one 36 amp-hour rechargeable battery.
- Design suitable for production at a cost well below $1,000 per unit.
- Weather-tight, animal resistant, and easily portable physical design.
- Operating environment -25C - +85C.

4.2.2. Power and Storage Requirements

The requirement of 14 day continuous operation on a 36 amp-hour battery leads to the following calculation for the average battery current:

$$\frac{36\,\text{amp hour}}{1}\cdot\frac{1\,\text{day}}{24\,\text{hours}}\cdot\frac{1}{14\,\text{days}}=107\,\text{milliamps}$$

.

The roughly 100mA average current at 6V nominal battery voltage indicates an average power consumption of just 600mW. This power limitation is particularly challenging because most existing hard drives require several hundred mA just to maintain the disk spinning. Therefore, it will be necessary to implement an audio cache system using solid-state memory (e.g., Flash memory) so that the disk need only be spun up when absolutely necessary. It is estimated that a 15 second spin up and write cycle every 10 minutes will meet the design goal.

It should also be noted that the battery size could be reduced (or battery life extended) if a supplemental power source such as photovoltaic panels (solar cells) or fuel cells were considered feasible.

The 14 day continuous recording capability will require considerable data storage capacity:

$$\frac{48\text{k samples}}{\text{sec}}\cdot\frac{3\,\text{byte}}{\text{samp}}\cdot\frac{60\,\text{sec}}{\text{min}}\cdot\frac{60\,\text{min}}{\text{hr}}\cdot\frac{24\,\text{hr}}{\text{dy}}\cdot\frac{14\,\text{dy}}{1}=162\,\text{gigabytes}$$

(where 1 gigabyte is 1,073,741,824 bytes). It is expected that the data storage requirement could be reduced by performing lossless data compression prior to placing the sound data onto the hard disk, so the actual storage necessary may be reduced to 80 GB or less. Drives of this size are becoming available in the sub-$100 range. Provision for more data channels, higher sampling rates, and other enhancements, would necessarily increase the required storage.

4.2.3. Other Features

Several additional features are anticipated for future development, as listed here.

- Integration of a GPS module for precise time-of-day and position determination. This feature could potentially allow acoustical beamforming and direction finding using time-aligned data from multiple independent sensors.
- Alternative power sources, including supplementary solar, fuel cell, and thermoelectric. Additional power for heating/cooling the system may also be needed for some applications.
- Streamlined packaging to allow NPS to send a system to a park superintendent with minimal training and configuration.
- High speed network access to allow data transfer from the recorder to a laptop or removable storage drive.
- Provision for additional data storage, such as meteorological observations.

It is expected that testing and evaluation of the prototype sound recording system will reveal the need for additional features and capabilities.

5. CONCLUSION

Since the 1970s the U.S. National Park System has been the host of many acoustical measurements. Although most studies have been in response to noise intrusions from tour aircraft, industrial operations, and motorized recreational vehicles, the development of new long-term recording devices and signal analysis procedures will extend the existing knowledge base regarding the sonic environment of the parks. Long-term data will be suitable for archiving, data analysis, park planning, and biological surveys. The availability of this data will allow correlation with other ecosystem measurements and trends.

6. ACKNOWLEDGEMENTS

Portions of this work were supported by Sanchez Industrial Design, Inc., and Advanced Acoustic Concepts, Inc. The authors also acknowledge the assistance of Bill Schmidt, Skip Ambrose, Shan Burson, and Bernie Krause.

7. REFERENCES

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

[2] Stanton, Robert (2000), "Director's Order #47: Soundscape Preservation and Noise Management," National Park Service.

[3] Krause, Bernard L. and Gage, Stuart (2003), "Testing Biophony as an Indicator of Habitat Fitness and Dynamics," Sequoia National Park (SEKI) Natural Soundscape Vital Signs Pilot Program Report, Wild Sanctuary, Inc., February, 2003. http://envirosonic.cevl.msu.edu/seki/Document.asp

[4] Langford, Nathaniel Pitt (1905), *The Discovery of Yellowstone Park*, Lincoln, NE: University of Nebraska Press, 1972 (from 1905 manuscript).

[5] Abbey, Edward (1968), *Desert Solitaire: A Season in the Wilderness*, New York: Simon and Schuster. Touchstone edition, 1990.

[6] Stegner, Wallace (1969), *The Sound of Mountain Water*, Garden City, NY: Doubleday & Co.

[7] Olson, Sigurd F. (1972), *Sigurd F. Olson's Wilderness Days*, New York: Alfred A. Knopf.

[8] Truman, Harry S. (1949), "Establishing an Airspace Reservation Over Certain Areas of the Superior National Forest in Minnesota," Executive Order 10092 (signed December 17, 1949), 14 FR 7681, December 22, 1949.

[9] Boundary Waters Canoe Area Wilderness Act 1978 Public Law 95-495, 92 Stat. 1649, Oct. 21, 1978.

[10] Wilderness Act, Public Law 88-577, 78 Stat. 890, 16 USC 1131, Sept. 3, 1964.

[11] Cox, E. F., Plagge, H. J., and Reed, J. W. (1954), "Meteorology Directs Where Blast Will Strike," Bulletin of the American Meteorological Society, vol. 35, no. 3.

[12] Shurcliff, William (1970), *S/S/T and Sonic Boom Handbook*, New York: Ballantine Books.

[13] Horwitch, Mel (1982), *Clipped Wings: the American SST Conflict*, Cambridge, MA: MIT Press.

[14] Welch, Bruce L. (1968), "SST: Coming Threat to Wilderness," National Parks Magazine, vol. 42, pp. 9-11.

[15] National Environmental Policy Act (1969), 42 USC 4321— 4370d; 83 Stat. 852, PL 91- 190, Jan. 1, 1970.

[16] United States Air Force (1996), "Sonic Boom," USAF Fact Sheet 96-03, March 1996.

[17] Nixon, Richard M. (1972), "Use of Off-road Vehicles on the Public Lands", Executive Order 11644, 37 FR 2877, 3 CFR, 1971-1975, February 8, 1972.

[18] Noise Control Act of 1972, Public Law 92-574, 86 Stat. 1234, 42 USC 4901, October 27, 1972.

[19] Ross, J.C. and Menge, C.W. (2002), "Draft Supplemental Technical Report on Noise: Winter Use Plan Final Supplemental EIS," HMMH Report 295860.400, October 2002.

[20] Foch, James D. (1992), "Bryce Canyon National Park and the Protection of Natural Quiet," Sound and Vibration, vol. 26(2), pp. 20-23.

[21] Grand Canyon National Park Enlargement Act (1972), Public Law 93-620, 88 Stat. 2091, 16 USC 228g, January 3, 1975.

[22] Hoffman, P., (2002), Deputy Assistant Secretary for Fish and Wildlife and Parks, U.S. Dept. of the Interior, statement before Senate Committee on Commerce, Science, and Transportation, October 3, 2002. www.nps.gov/legal/testimony/107th/overflit.htm

[23] Brumbaugh, D. S. (1985), "A Report on the Analysis of the Effect of Helicopter Vibrations on the Pt. Sublime Anasazi Site, Grand Canyon National Park," Flagstaff, AZ: NAU Department of Geology, 1985.

[24] National Park Service (1995), "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.

[25] National Transportation Safety Board (1987), "Midair Collision Over Grand Canyon National Park, June 18, 1986," Report NTSB/AAR-87/03.

[26] "Aircraft Overflights in National Parks," Public Law 100-91 (101 Stat. 674), August 18, 1987.

[27] Anderson, G. S., and Horonjeff, R. D. (1992), "Aircraft Overflight Study. Effect of Aircraft Altitude upon Sound Levels at the Ground," Report NPOA-91-4, HMMH-290940.02 (NTIS Number: PB93-144194), March 1992.

[28] Miller, N. P., Sanchez, G., and Anderson, G. S. (1995), "Aircraft Management Studies: Selecting a Simplified Method for Acoustic Sampling of Aircraft and Background Sound Levels in National Parks," Report NPOA-95-1, HMMH-290940.24 (NTIS Number: PB95-263307), July 1995.

[29] Tabachnick, B. G., Howe, R. R., and Fidell, S. (1992), "Estimation of Aircraft Overflight Exposure in National Parks and Forest Service Wildernesses," Report NPOA-92-1, BBN-7259 (NTIS Number: PB93-144293), August 1992.

[30] Anderson, G. S., Horonjeff, R. D., Menge, C. W., Miller, N. P., Robert, W. E., Rossano, C., Sanchez, G., Baumgartner, R. M., and McDonald, C., "Dose-Response Relationships Derived from Data Collected at Grand Canyon, Haleakala and Hawaii Volcanoes National Parks," Report NPOA-93-6, HMMH-290940.14 (NTIS Number: PB94-151941), October 1993.

[31] McKechnie, A. M., and Gladwin, D. N. (1993), "Aircraft Overflight Effects on Wildlife Resources," Report NPOA-93-8, HMMH-290940.22 (NTIS Number: PB94-149994), November 1993.

[32] Hanson, C. E., King, K. W., Eagan, M. E., and Horonjeff, R. D. (1991), "Aircraft Noise Effects on Cultural Resources: Review of Technical Literature," Report NPOA-91-3, HMMH-290940.04-1 (NTIS Number: PB93-205318), September 1991.

[33] Harris, Miller, Miller and Hanson, Inc. (1994), "Survey of National Park Service Managers Related to Aircraft Overflying National Parks," Report NPOA-93-7, HMMH-290940.17 (NTIS Number: PB95-105896), May 1994.

[34] Harris, Miller, Miller and Hanson, Inc. (1994), "Aircraft Management Studies Air Tour Passengers Survey," Report NPOA-94-1, HMMH-290940.15 (NTIS Number: PB95-104014), May 1994.

[35] Federal Aviation Administration (FAA), Special Federal Aviation Regulation (SFAR) 50-2, September, 1988.

[36] Fidell, S., Pearsons, K., and Sneddon, M. (1994), "Evaluation of the Effectiveness of SFAR 50-2 in Restoring Natural Quiet to Grand Canyon National Park," Report NPOA-93-1, BBN-7197 (NTIS Number: PB95-195202), June 1994.

[37] "National Parks Air Tour Management Act of 2000," Public Law 106-181 (114 Stat. 186), April 5, 2000.

[38] Fleming, Gregg G., Roof, Christopher J., and Read, David R. (1998), "Draft Guidelines for the

Measurement and Assessment of Low-Level Ambient Noise," U.S. Dept. of Transportion, John A. Volpe National Transportation Systems Center, Report DTS-34-FA865-LR1, March 9, 1998.

[39] Ambrose, Skip (2003), "Air Tour Management Planning and Acoustic Data Collection for National Parks," National Park Service Natural Sounds Program, April 2003.

[40] Ambrose, Skip (2004), private communication.

[41] Burson, Shan (2005), private communication.

[42] Sanchez, Gonzalo (2005), Sanchez Industrial Design, Inc., Middleton, WI, USA. Private communication.

[43] Bregman, A. S. (1990), *Auditory Scene Analysis: The Perceptual Organization of Sound*, Cambridge, MA: MIT Press.

[44] Gygi , Brian (2001), "Factors in the Identification of Environmental Sounds," Ph.D. dissertation, Department of Psychology, Indiana University, July, 2001.

[45] Gregoire, B. J. and Maher, R. C. (2005), "Harmonic Envelope Detection and Amplitude Estimation Using Map Seeking Circuits," Proc. IEEE International Conference on Electro Information Technology (EIT2005), Lincoln, NE, May, 2005.

[46] Sanchez, Gonzalo (2003), *Preliminary Datasheet: Model PADR-100 Portable Audio Data Recorder*, Sanchez Industrial Design, Inc., Middleton, WI, USA.

[47] Miller, Nicholas (2001), "A Proposal for Acoustic Data Collection in Parks and Wilderness Areas," Proc. 2001 Int. Congress and Exposition on Noise Control Eng. (Inter-noise), The Hague, The Netherlands, August 2001.

[48] Miller, Nicholas, and Menge, Christopher (2001), "Status of the I-INCE Initiative on Recreational Noise and Progress on Quantifying Noise Intrusions in Parks," Proc. NOISE-CON 2001, Portland, ME, October 2001.