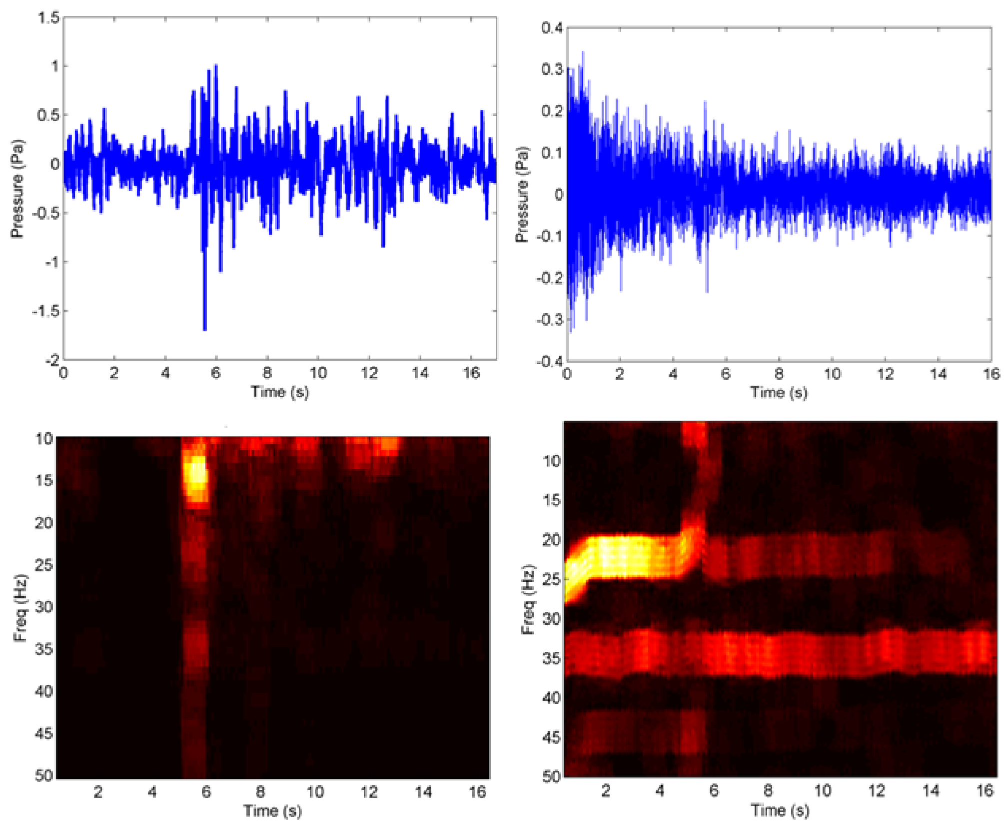
ERDC TR-12-12

US Army Corps of Engineers®
Engineer Research and Development Center

Data Preparation Procedures for the ERDC High-Energy Large-Scale Blast Sound Propagation Experiment

Daniel P. Valente, Lauren Ronsse, Roger D. Serwy, Jesse Barr, Keran Claffey, Michael J. White, and Michelle E. Swearingen

September 2012



Approved for public release; distribution is unlimited.

Engineer Research and Development Center

ERDC TR-12-12
September 2012

Data Preparation Procedures for the ERDC High-Energy Large-Scale Blast Sound Propagation Experiment

Daniel P. Valente, Lauren Ronsse, Roger D. Serwy, Jesse Barr,
Michael J. White, and Michelle E. Swearingen

Construction Engineering Research Laboratory (CERL)
US Army Engineer Research and Development Center
2902 Newmark Dr.
Champaign, IL 61822-1076

Keran Claffey

Cold Regions Research Laboratory (CRREL)
US Army Engineer Research and Development Center
72 Lyme Rd.
Hanover, NH 03755-1290

Final Report

Prepared for Headquarters, US Army Corps of Engineers
Washington, DC 20314-1000

Under Work Unit C43279

Abstract

This work generated acoustic data for nearly 28,000 blast events in a variety of weather conditions, and for associated meteorological conditions. Methods were developed to reduce and quality-check those data to create a dataset of over 18,000 curated blast waveforms, associated acoustic metrics, and corresponding meteorological data. This dataset provides a resource unique to the US Army Engineer Research and Development Center (ERDC) that may be used to assess the statistical variability of blasts noise as a function of distance, climate, and meteorology, and to investigate the definition of acoustic propagation classes for blast noise.

Table of Contents

Preface

This study was conducted for the US Army Public Health Command (USAPHC) as part of the Operational Noise Discrete Events: Physics Models Environmental Quality Technology (EQT) program under Project D048, “Industrial Operations & Pollution Control Technology,” Work Unit C43279, “Blast Propagation Data Analysis.” The technical monitor was Catherine Stewart, MCHB-IP-EON.

The work was performed by the Ecological Processes Branch (CN-N) of the Installations Division (CN), , US Army Engineer Research and Development Center – Construction Engineering Research Laboratory (ERDC-CERL), and the Signature Physics Branch (RR-D), Cold Regions Research and Engineering Laboratory (CRREL). The CERL Principal Investigator was Larry L. Pater. William Meyer is Chief, CEERD-CN-N, and Dr. John Bandy is the Chief, CEERD-CN. The Deputy Director of ERDC-CERL was Dr. Kirankumar V. Topudurti and the Director was Dr. Ilker R. Adiguzel.

The Commander and Executive Director of ERDC is COL Kevin J. Wilson, and the Director of ERDC is Dr. Jeffery P. Holland.

Unit Conversion Factors

Multiply	By	To Obtain
atmosphere (standard)	101.325	kilopascals
Bars	100	kilopascals
degrees (angle)	0.01745329	radians
degrees Fahrenheit	(F-32)/1.8	degrees Celsius
Feet	0.3048	meters
Inches	0.0254	meters
miles (nautical)	1,852	meters
miles (US statute)	1,609.347	meters
miles per hour	0.44704	meters per second
pounds (force) per square foot	47.88026	pascals
pounds (force) per square inch	6.894757	kilopascals
pounds (mass)	0.45359237	kilograms

1 Introduction

1.1 Background

Sound generated from large explosions (i.e., blast noise) is one of the most common sources of noise around Army installations. A blast initiates with a rapid exothermic chemical reaction and an increase in pressure so drastic that a discontinuous pressure boundary arises in the explosive (Cooper 1996). From an acoustics perspective, this high-energy pressure pulse, more commonly referred to as the *shock*, imbues blasts with unique acoustical properties. Not only does the shock drive the exothermic reaction by heating the remaining explosive, but it propagates into the surrounding fluid. To an observer a safe distance away, it is recognized as the “bang” of blast noise. This “bang” is impulsive and rich in low-frequency acoustical energy. Compared to other sources of noise, e.g., vehicles or power generators, the high energies contained within the blast at these low frequencies allow for minimal atmospheric absorption of the sound and only small losses from interactions with terrain; in fact, blast noise can propagate to distances in excess of 20km.

It is not surprising, then, that weather significantly affects blast sound propagation. For moderate to long-distances (e.g., several kilometers), received sound level can vary by as much as 50 dB due solely to changes in the vertical effective sound speed profile (Schomer 2001). This effective sound speed profile is strongly dependent on the meteorological parameters of wind speed, wind direction, and temperature. Under certain atmospheric conditions, for example, an observer 10km south of a blast site may be barely able to hear a particular blast event, whereas someone the same distance north of the site will experience the sound as quite loud. This variability makes it notoriously difficult to predict blast noise levels with a high degree of accuracy.

From an Army operations standpoint, there is a need to better understand long-range sound propagation of these signals for a number of purposes: acoustical tracking and source identification on the battlefield, development and validation of sound prediction models, and avoidance of adverse community reaction to noise that can compromise training and testing missions. Detailed, accurate meteorological and acoustical experimental data, for a large range of meteorological conditions that occur throughout the day and year and in various climate types, are required to improve our understanding of the statistics of long-distance received sound level. These data can then guide definition of propagation classes—meteorology-based categories with unique acoustic propagation statistics. Such classes are essential to test analytical algorithms and develop noise impact management decision guidance. Requirements for this research are based in Army Environmental Requirements and Technology Assessments (AERTA) requirement 2.4.f, DoD Instruction 4715.13, and Army Regulation (AR) 200-1.

1.2 Objectives

The overall objective of this project was to measure an extensive set of high-energy impulsive pressure waveforms and simultaneous meteorological data at distances up to 16 km. Using these data, meteorological and ground conditions were to be partitioned into a set of propagation classes with associated statistics.

The specific objectives of this initial stage of research were to:

1. Develop and detail procedures to be used to reduce the acoustical data into a readily-usable format
2. Clean the data prior to analysis
3. Process the resulting waveforms into pertinent acoustic metrics, namely the unweighted peak level (L_{pk}) and the sound exposure level (SEL)
4. Develop and document a methodology to reduce the meteorological data into a readily-usable format.

1.3 Approach

To prepare the data for analysis, raw acoustic recordings were pre-processed and reduced to a smaller, more manageable dataset. To this end, all of the blast waveforms were windowed and several metrics were calculated to describe these signals. The waveforms were saved into individual files, along with associated metadata and metrics. The pre-processed waveforms were then subject to a rigorous data cleaning procedure in preparation for subsequent statistical analysis.

1.4 Mode of technology transfer

This report will be provided directly to the Operational Noise Program of USAPHC (the Army technical transfer agent for and primary user of military blast noise technology) and to other known users, in addition to being made accessible through the World Wide Web (WWW) at URLs:

http://www.cecer.army.mil

http://libweb.erdc.usace.army.mil

2 Reduction of the Acoustic Data

In the ERDC High-Energy Large-Scale Blast Sound Propagation experiment, blast noise measurements were taken at two Army installations (a desert climate and a temperate climate) at two times of year (summer and winter). In each experiment, charges of plastic explosive composition C-4 were detonated approximately every 20 minutes for 6-8 hours each day for 10 days. Acoustic sensors were positioned at distances ranging from 4m to 16km, at three different angles from the blast source location (referred to as “measurement arms” or “lines”). The goal of the reduction process was to obtain an individual blast waveform for each blast recorded by each sensor in the experiment. Further details of the experiment are given in Pater et al. (2012).

Data acquisition for the sensors closest to the source (4m and 125m) was directly triggered by the blast event and recorded on a Yokogawa digital recording oscilloscope. Data at the remote sites, however, could not be easily triggered and associated to a specific blast event. Data were therefore recorded continuously at these sites in files that reached up to 2 GB each day. To reduce computational overhead and eliminate data that did not contain relevant blasts, the data were reduced using the blast times defined by initiation of the trigger on the sensors at 4m. Based on this trigger, time windows for expected blast arrivals were constructed for the remote sites and these windows were extracted from the day-long continuous recording. A collection of MATLAB* scripts and functions facilitated the data reduction process. All MATLAB code used in the reduction process can be obtained by request from the authors of this report.†

The reduction process is outlined as:

1. The data were placed into a pre-defined directory structure.
2. Time windows of interest were identified, waveforms were extracted into individual files, and metadata associated with each waveform were stored.
3. Calibration values were calculated and appended to the metadata file.

* MATLAB (**MAT**rix **LAB**oratory) is a numerical computing environment and fourth-generation programming language. Developed by The MathWorks, Inc., Natick MA (http://www.mathworks.com/).

† Authors may be contacted via the street address provided on the first page of this report.

This resulted in a preliminary dataset that was then subjected to a rigorous quality control procedure (described in Chapter 3).

2.1 The assumed directory structure

The MATLAB scripts written to facilitate data reduction assume a specific directory structure for the raw data. Since the directory path is used to label the data, it must adhere to this directory structure to process properly. With these scripts, all of the recorded data is assumed to follow the directory structure below:

...\Recorded\YYYY-MM-DD\STATION\...

The four data sets are:

- NM07 – White Sands Missile Range, NM 2007 (summer)
- NM08 – White Sands Missile Range, NM 2008 (winter)
- MO08 – Fort Leonard Wood, MO 2008 (summer)
- MO09 – Fort Leonard Wood, MO 2009 (winter).

The sensor locations (i.e., station names) are defined by a letter, indicating the measurement arm (A, B, or C), and a distance in kilometers from the source point. Additional sites were located near the source point (G0), and are labeled Y-#, indicating the channel number on the Yokogawa recording oscilloscope, for example, the folder F:\NM08\Recorded\2008-01-23\A12\ stores all files from the White Sands 2008 experiment that took place on 23 January 2008 and were recorded on the A line at 12km. Details regarding actual locations and instrumentation are given in Pater et al. (2012).

2.2 Time windowing and storing waveforms

The recorders at the remote sites (A, B, and C lines) ran continuously during the experiment periods, recording up to 11 hours of data in a single file of up to 2 GB in size. To streamline the data reduction process and to ensure that only blast signals were analyzed, synchronized timestamps were used to find the blasts, and the time-windowed waveforms were saved in separate files. The files containing individual waveforms were then quality-checked to ensure that the signal of interest (the blast) was actually captured, i.e., that the data were not buried in background noise.

Each recording device (the Yokogowa digital oscilloscope or the Rion data recorders) recorded an Inter-Range Instrumentation Group mod B (IRIG-B) timecode signal on one of its available channels. The IRIG-B signal is deterministic and well-defined (IRIG Standard 200-98), containing “ticks” that occur once per second, with a sampling rate of 1 kHz and using amplitude modulation to encode binary time data on a per-cycle basis. The timecode generators were synchronized via global positioning system (GPS), thereby ensuring that a particular IRIG-B tick on one sensor occurred at the same instant as that same tick on another.

The shot times were obtained from the Yokogawa data recorders. Each blast was captured in a separate file, and each file consisted of 5 seconds of data. The first IRIG-B timecode decoded from a particular blast signal at one of the 4 m sensors was used as the time label for that blast, and the IRIG-B tick was used as the 0-second time mark. By defining the starting point as the first IRIG-B tick recorded by the Yokogawa, all the data can be referenced to a well-defined starting point, almost down to the exact sample recorded. Because there was 1 second of pre-trigger recorded, the blast onset will be randomly distributed between zero and 1 second relative to the IRIG-B tick. The blast onset was not chosen as the zero-point because it is not well-defined. Each analyst may have a different definition, such as peak pressure, –20dB from peak pressure, etc.

The reduction process searches for a given IRIG-B timecode in each file. Next, a time-windowed portion of the waveform is saved. The window is defined as delays from the IRIG-B time. The delays are in seconds, and can be negative. Specifically, given an IRIG-B time A, the window is:

$$[A + L, A + R]$$

where:

L = the start time delay
R = the end time delay.

L and R are defined as:

$$L = d/370 - 2$$
$$R = d/330 + 11$$

where:

d = the distance from the source point to the receiver point
370 m/s = an unrealistically high sound speed
330 m/s = an unrealistically low sound speed.

A typical sound speed in air is 340 m/s at 15 °C when not affected by other meteorological conditions. The additional 2 s delay on L and the 11 s buffer on R ensure that the window encompasses the entire blast signal, allowing for high variability in instantaneous sound speeds as the blast propagated through the atmosphere, and multiple arrivals from various atmospheric paths.

The raw 16-bit data within this window is saved from the original data recording, along with the scaling and offset factors provided by the data recorder. This ensures data integrity. These files are given the `extension.blast`, but are essentially `MATLAB.mat` files containing the relevant information.

A metadata file is also created at this time, and contains information about the windowed waveform. This file contains the IRIG-B timecode of the original blast time, the first IRIG-B timecode found in the windowed waveform range, the sample of the IRIG-B tick, the sampling rate, number of samples, and placeholder calibration information (offset, calibration constant, units). Calibration information used to transform the waveform from Volts into Pascals is obtained in a later step.

For every station and channel, for each blast, the output of the reduction process is:

```
reduction_path\IrigB-time\Station-Channel.blast
reduction_path\IrigB-time\Station-Channel.meta
```

For example, the files:

```
F:\Reduced\08-001-11-20-20\A12-1.blast
F:\Reduced\08-001-11-20-20\A12-1.meta
```

correspond to the blast waveform and metadata associated with the blast that occurred at IRIG-B time 08-001-11-20-20 (1 January 2008 11:20:20), on Channel 1 of the 12km sensor on Line A.

2.3 Calibration

Once the raw waveforms are windowed and saved as separate files, they must be calibrated. Applying the sensor-specific calibration constant converts the raw data from Volts into Pascals, the physical quantity of interest. The process of determining the appropriate file set to use for calculating the calibration constants is similar to the process described for finding the waveforms. The required directory structure is the same, but the files do not need to be partitioned into individual waveforms. Calibration files contain 30 seconds to 1 minute of sinusoidal data in comparison to the 11 hours of the continuous daily recordings. The files are downsampled and the average value of the sine wave is found. This information, combined with the known frequency and amplitude of the calibrator, enables calculation of the calibration constant. To produce calibrated waveforms from the data, these microphone-specific values are multiplied to the appropriate waveforms.

All the data recorders record voltages. The calibration rules are linear. Given the original voltage waveform x, the calibration rule to convert the waveform from Volts to Pascals is:

$$y = mx + b,$$

where:

m = the calibration scaling factor
b = the calibration offset.

In other words:

```
pressure_signal = cal_scale * voltage_signal + cal_offset
```

Appendix A to this report contains detailed instructions for the running the reduction and calibration procedures.

3 Data Cleaning/Quality Control

The data reduction process resulted in a set of 28,821 blast waveforms and associated metadata files. These preliminary data were then subjected to a rigorous data cleaning procedure to remove waveforms that did not contain blast signatures and to discover incorrect calibration values. The absence of a usable blast signature can be attributed to either: (1) equipment failures, (2) failed explosions, i.e., "duds," (3) blast levels below the equipment noise level, (4) blasts below the ambient noise level, or (5) blasts that did not reach the sensor due to atmospheric conditions. Incorrect calibration values can be attributed to human error in the data reduction process.

3.1 QC Level 1: Human curation

The first level quality control procedure consisted of a labor-intensive human curation step. A MATLAB function, `classifyLRPE`, was written to expedite the curation process. The `classifyLRPE` function is an interactive function that reads in an index of blast waveforms, steps through each file in the index, and allows users to append a classification of "blast" or "no blast" to the blast metadata file. As the user steps through each file in the index, the waveform is automatically displayed. At this point, the user can choose to classify the waveform (blast/no blast), to view the multitaper spectrogram of the waveform (Percival and Walden 1993), or to listen to the blast. For any option the user selects, the possibility of repeating any of the other options is retained until the user is comfortable declaring whether a blast is present in the waveform. If the user determines that a blast is present, they are prompted to click on the waveform or spectrogram at a point immediately before the blast occurs. These "approximate blast start times" are used to define a narrow window surrounding the blast event and are saved in the metadata file; the times facilitate the subsequent analysis of acoustic quantities. Once the classification has been made, the filename is annotated with the associated classification and the next blast file is presented to the user.

The choice to allow users to examine plots of the waveform and/or spectrogram, in addition to listening, allows for faster classification and better recognition of blasts that are present in the waveform, but that are inaudible or obscured by noise. The larger goal of the study is to obtain the true distribution of received levels as accurately as possible, and this distribution undoubtedly contains inaudible blasts, especially at extreme distances from the source. Listening is limited by the capabilities of the sound reproduction system, the Signal-to-Noise Ratio (SNR) of the blast, and the hearing ability of the listener. The participating listeners exclusively listened through headphones, and because of the high energy at low frequencies typical of blasts, it is difficult for many headphones to adequately reproduce blast noise. Furthermore, other noise sources (wind, vehicles, aircraft, electrical noise etc.) can mask the presence of a blast to a human listener.

For sensors that are within 4km of the blast, it is typically possible to determine the presence of a blast in the waveform based on visual inspection of the waveform alone; the amplitude at these distances is usually greater than 100 dB peaks with SNRs in excess of 20dB. Therefore, the close distances can be quickly identified and confirmed as containing blasts, if present. Beyond this distance (and for some cases at short distances with high ambient noise) the multitaper spectrogram provides a useful tool for blast recognition, even in cases where a human listener cannot hear the blast. The multitaper spectral estimate provides an unbiased estimate of the spectrogram with lower variance than other tapering methods, therefore reducing unwanted noise. A blast is an impulsive event with energy concentrated at very low frequencies (typically 10-100Hz), and appears as an abrupt, but transient rise in broadband energy in the spectrogram. The multitaper method has the additional "benefit" (for the purpose of blast classification) of slightly smearing the blast event in the time frequency plane, so that an impulsive sound manifests as a rectangle of high amplitude (Figure 1). The rectangular feature is most visible in the 10-100Hz band. This allows for a quick visual determination of the presence of a blast in the waveform, even with high amplitude contamination from other noise sources (Figure 2).

To begin the curation process, a master index of all waveform files in the dataset was compiled, and then divided into 15 sub-indexes of approximately 2000 files each. These indexes were then assigned to four "novice" listeners. The four novice listeners were each given the same instructions: to go through each blast in their assigned index and use any or all of the three methods provided by `classifyLRPE` to determine if a blast is present, then use `classifyLRPE` to annotate the waveform.

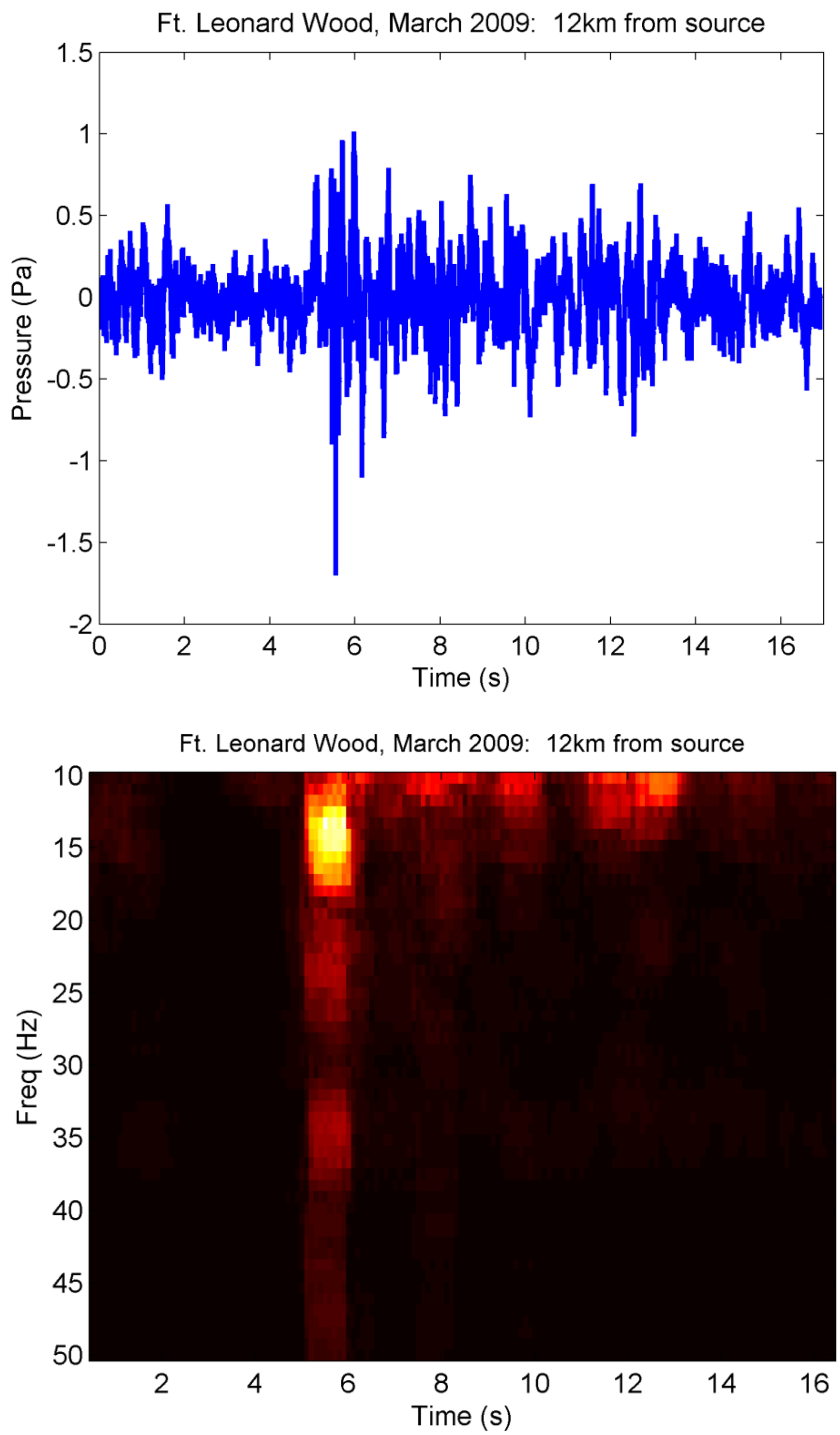


Figure 1. Illustration of blast signature in the multitaper spectrogram at low frequencies: waveform (top panel) and spectrogram in the 10-50 Hz band (bottom panel).

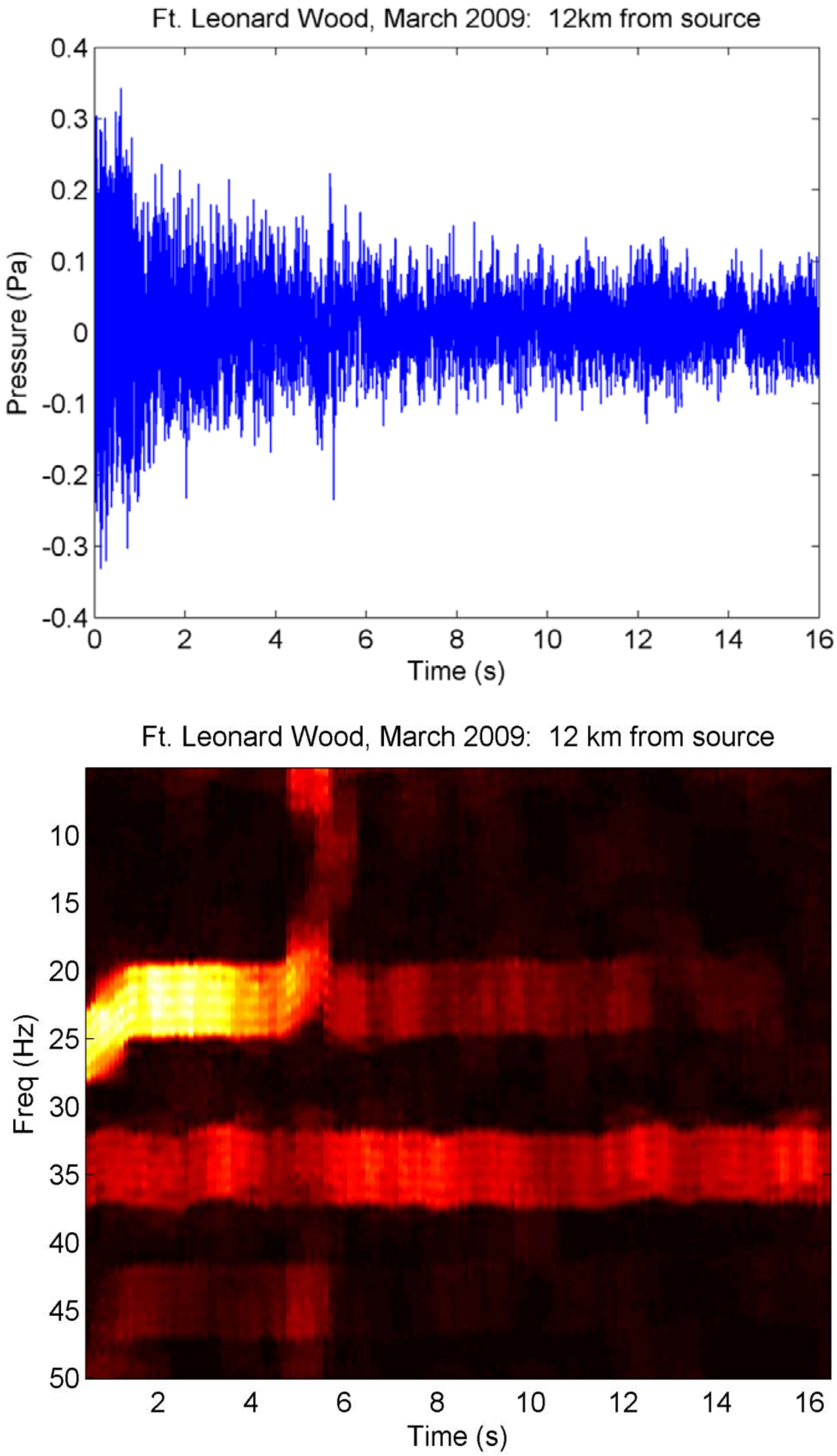


Figure 2. Example demonstrating that a blast is visible in the multitaper spectrogram in the presence of high amplitude noise contamination. The blast signature appears on the spectrogram (right panel) as the vertical rectangular feature at approximately 5 seconds. Tonal low-frequency noise is shown in the horizontal rectangles.

As each person finished classifying the 2000 files in an index, they were assigned another. The four individuals completed a different number of indexes due to the speed at which each were able to classify. Furthermore, although they were given the same instructions, there was variability in the strategies that the listeners employed in determining a blast, which range from listening to a majority of the files, to exclusively using the spectrogram.

3.2 QC Level 2: Expert confirmation

Due of the size of the dataset, the variability in classification strategies among the four participants, and errors noticed in the QC Level 1, a second QC step was enforced. The errors were first noticed in the blast start times as the data analysis procedure began, and further investigation uncovered that other instructions were not followed as given. For instance, it was noticed that some waveforms with high background noise, no audible blast, and no blast signature on the spectrogram were classified as blasts.

For this step, analysis was restricted to the files that were identified as blasts in QC Level 1. Using a modification of `classifyLRPE` named `relisten`, and following the same instructions given to the four novice listeners, two expert listeners reexamined each waveform. These were the same individuals who were to perform final analysis on the data (Dan Valente and Lauren Ronsse). The expert listeners used a somewhat conservative approach to classification and discarded any file that they were not absolutely sure contained a blast. During this process, incorrect blast start times were also corrected. Finally, the dataset was restricted to blasts originating from charge sizes of approximately 1.25 lbs (1.2 – 1.4 lbs), which was a necessary step for subsequent statistical analyses.

3.3 Calculation of basic acoustic quantities

The second level QC procedure resulted in a list of waveforms with high confidence of containing a blast signature. Basic acoustic quantities were then calculated for each blast in the dataset, using the function `calcBasicQuantities`. These quantities include the peak sound pressure level for each blast (L_{pk}), two separate SNRs, and the SEL. Also associated with each blast is the filename, the local time of occurrence (as measured at ground zero), the station name, the station distance, and the charge size.

The original blast waveforms, resulting from the data reduction process, are approximately 15 seconds long. To calculate accurate acoustic quantities associated with the blasts, and not of unwanted noise sources, the approximate blast start times found in the above QC procedure were used to window the data. Slightly different window lengths were used for the sensors located at 4m than were used for the other sensor distances. Blast waveforms near the source experience very little atmospheric spreading, so a 1-second window was chosen to capture the blast. For stations that were distances greater than 4m, a 3 second window was decided on to capture sufficient reverberation while still minimizing contamination from other noise sources. In the case of the 1-second window, the window was defined as extending from 100ms prior to the approximate start time to 900ms after the approximate start time. In the 3-second case, the window extended from 1 second before the approximate start time to 2 second after.

Next, the location of the blasts within the defined windows was standardized. To do this, the peak of the waveform within the above window was defined as a landmark, and the window was fixed around this point (that is, instead of being fixed around the approximate start time). Within this new window, the waveform was mean-centered, and the unweighted peak level and unweighted SELs were calculated.

The importance of mean-centering the waveform (i.e., subtracting the DC [direct current] level) is a subtle, yet necessary step in the data cleaning procedure. The DC value at any point is a sum of the DC value due to the electronics, the ambient barometric pressure, and the actual acoustic process being measured. The DC value of the waveforms at each sensor tracked very closely with temperature (for an example see Figure 3), implying that either the microphones, preamps, or other electronics exhibited temperature dependence. Failure to account for this effect would result in incorrect acoustic values and an artifactual dependence on local ambient temperature.

Finally, two SNRs were examined for their utility to characterize the blast with respect to ambient levels. The first was the crest factor, or the peak-to-rms value of the waveform. The second was the signal variance to the noise variance. The relationship between the different SNR calculations was approximately linear; therefore, only the peak-to-rms SNR was used in the subsequent QC Level 3.

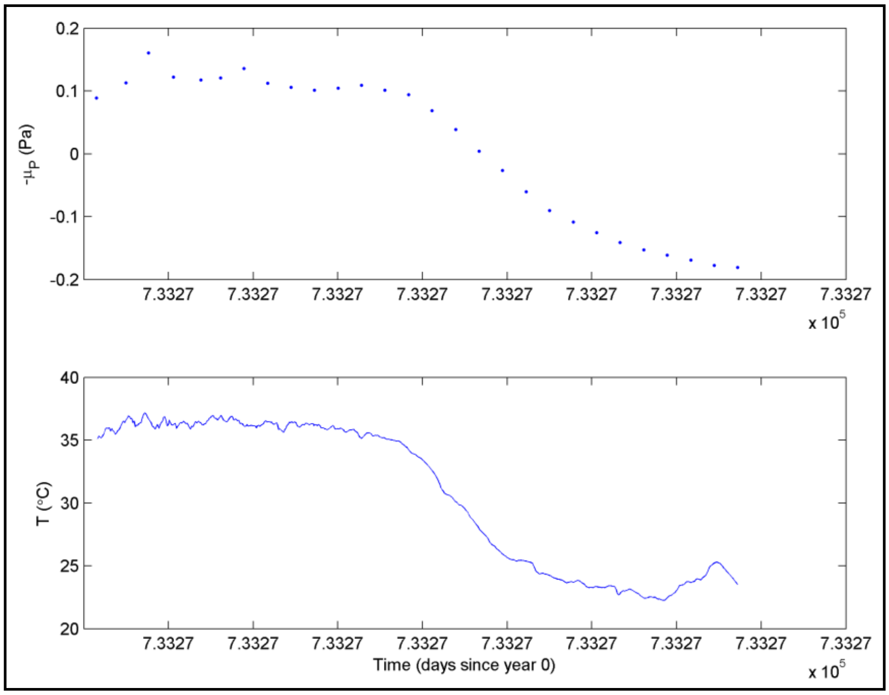


Figure 3. DC value for a measured waveform (Top panel) and ambient temperature (bottom panel) as a function of time for a single sensor (at 1km) in the NM07 experiment. DC value tracks closely with ambient temperature, demonstrating the importance of removing it from the signal before analysis of acoustic quantities.

3.4 QC Level 3: Anomaly detection and removal/repair

On calculating these preliminary acoustics quantities, a final quality control step was employed to verify that calibrations were correct, that the correct times had been selected, and that any equipment problems were realized and accounted for. To do this, the peak level vs. SNR and peak level vs. SELs were examined. Working under the assumption that peak levels and SELs should follow a relatively smooth distribution, yet have variable range of SNRs, outliers were singled out for further analysis. Typically, outliers were anomalous signals that were either duds, had unnecessary electrical noise (therefore suggesting an equipment problem), or experienced calibration problems that had escaped the data reduction process and the first two QC levels. In addition to outliers, several waveforms within the distributions at each sensor were examined for reasonable consistency.

The procedure was repeated for each of the four experiment sites, resulting in a final set of 18,251 blast files. Table 1 lists the number of blasts measured at each sensor for each experiment, along with total blasts of size ~1.25lbs that were attempted.

Table 1. Number of shots received at each sensor after QC procedure has been applied. The number in parentheses next to the experiment label is the number of attempted shots of size ~1.25 lbs.

Distance (km)	M008 (224)			M009 (207)			NM07 (231)			NM08 (227)		
	A	B	C	A	B	C	A	B	C	A	B	C
0.004	214	216	216	194	195	198	227	222	203	222	222	222
0.125	216	216	216	198	197	196	218	227	227	213	222	221
1	219	203	199	91	202	202	231	231	231	182	225	206
2	216	217	218	170	197	200	231	202	228	215	225	225
4	209	208	216	129	162	196	230	203	229	207	224	208
8	179	208	217	108	172	171	223	181	211	203	160	192
12	149	148	209	89	144	153	204	159	191	159	196	184
14	205	–	–	125	–	–	–	–	–	–	–	–
15	–	–	–	–	–	–	–	–	149	–	–	143
16	–	148	189	–	152	141	–	113	–	–	171	–

The documentation of each function in `analyzeLRPE`, the driver script for the quality control/data cleaning procedure, gives more details on how the procedure is implemented. The necessary functions can be obtained by request from the authors of this report.

4 Meteorological Data

Data was collected on two separate systems:

1. A tower-mounted HOBO® data logger system to measure the atmospheric properties up to a height of 15 m
2. A tethersonde system to measure the atmospheric properties up to approximately 2km.

Software packages provided by the equipment manufacturers were used to convert the raw data from a proprietary format into Microsoft® Excel® spreadsheets. All meteorological data from the experiments can be obtained by request from the authors of this report.

The HOBO data were further compiled into four master spreadsheets, one corresponding to each experiment. Each record in these spreadsheets corresponded to a sample time and each column contained the data from the sensors on each of the weather towers. This structure (with each record as a time sample) allows for a straightforward comparison to the blast times provided during the reduction process.

5 Conclusion

This work generated acoustic data for nearly 28,000 blast events in a variety of weather conditions, and for associated meteorological conditions. Methods were developed to reduce and quality-check those data to create a dataset of over 18,000 curated blast waveforms, associated acoustic metrics, and corresponding meteorological data. Specifically:

1. Procedures to reduce the acoustical data into a readily-usable format were developed and documented.
2. Data were cleaned prior to analysis.
3. The resulting waveforms were processed into pertinent acoustic metrics, namely the unweighted peak level (L_{pk}) and the SEL.
4. A methodology to reduce the meteorological data into a readily-usable format was developed and documented.

This dataset provides a resource unique to the US Army Engineer Research and Development Center (ERDC) that may be used to assess the statistical variability of blasts noise as a function of distance, climate, and meteorology. This dataset provides, for the first time, a means by which to characterize the statistical variability of received level as a function of distance under different weather conditions. Furthermore, it allows for investigation into the definition of acoustic propagation classes for blast noise.

In-depth analysis of the acoustic data, how it segregates into propagation classes, and how it correlates to the meteorology will be the subject of subsequent reports.

The MATLAB routines written for reduction and quality control of the acoustic data are available by request, and can be used with little modification to reduce and clean data from experiments of similar design.

Acronyms and Abbreviations

Term	Definition
AERTA	Army Environmental Requirements and Technology Assessments
AR	Army Regulation
CEERD	US Army Corps of Engineers, Engineer Research and Development Center
CERL	Construction Engineering Research Laboratory
CRREL	Cold Regions Research and Engineering Laboratory
CSV	Comma Separated Values
DC	Direct Current
EQT	Environmental Quality Technology
ERDC	Engineer Research and Development Center
GB	Gigabyte
GPS	Global Positioning System
IRIG	Inter-Range Instrumentation Group
IRIG-B	Inter Range Instrumentation Group mod B (IRIG-B)
MATLAB	MATrix LABoratory
OMB	Office of Management and Budget
QC	Quality Control
SEL	Sound Exposure Level
SF	Standard Form
SNR	Signal-to-Noise Ratio
TR	Technical Report
USAPHC	US Army Public Health Command
WWW	World Wide Web

References

Cooper, Paul W. 1996. Explosives engineering. New York: Wiley-VHC.

Timing Committee, Telecommunications and Timing Group, Range Commanders Council. 1998. IRIG serial time code formats. IRIG Standard 200-98. 1998. US Army White Sands Missile Range, NM: Secretariat Range Commanders Council.

Pater, Larry, Michael J. White, Donald Albert, Michelle E. Swearingen, D. Keith Wilson, Edward T. Nykaza, Bruce Macallister, Jeff Mifflin, Bonnie Jones, Dan Valente, and Sarah Nemeth. 2012. High energy large scale blast sound propagation experiments. Draft Engineer Research and Development Center (ERDC) Technical Report. Champaign, IL: Engineer Research and Development Center, Construction Engineering Research Laboratory (ERDC-CERL).

Percival, Donald B., and Andrew T. Walden. 1993. Spectral analysis for physical applications: Multitaper and conventional univariate techniques. Cambridge, UK: Cambridge University Press.

Schomer P. D. 2001. A statistical description of blast sound propagation. Noise Control Engineering Journal 49(2):79-87.

Appendix A: Detailed Description of the Data Reduction Process

The data reduction process is implemented as a series of MATLAB functions, driven by the `reduce_cell` script. The `reduce_cell` script uses MATLAB's cell mode feature, allowing sub-blocks to be executed as the user sees fit. The reader is referred to the MATLAB cell mode documentation for more information.

1. The functions that are called by `reduce_cell` for the reduction phase, with a brief description are:

   `indexFiles`, which creates a comma-separated values (CSV) file of all the files to reduce
   `redStep1`, which finds IRIG-B times from Yokogawa files
   `redStep2`, which catalogs the IRIG-B ranges of all the files to reduce
   `redStep3`, which reduces the waveforms given the blast times and rules for reduction.

2. The functions called by `reduce_cell` for calibration are:

   `calStep1`, which preprocesses calibration recordings. Downsamples and profiles.
   `calStep2`, which creates a CSV of all the calibration data found.
   `validateCalRules`, which validates that the user-generated calibration rules are correct.
   `calStep3`, which adds the calibration rules to the meta files.
   `verifyReduction`, which creates a CSV of all the channels reduced for each blast. Identifies missing channels and missing calibrations.

3. Source code for all files is available by request from the authors.
4. The driver scripts adhere to a reasonably strict nomenclature. For the scripts that require a user-generated or user-modified CSV file, the file name should be prefaced by `USER_`. When a function generates a CSV file, the file name is prefaced by a `GEN_`.

Steps for Reduction

The following step-by-step instructions detail how a user would use the functions in `reduce_cell` to reduce the data.

1. Define directory locations for the storage of the raw data, the reduced data, and the helper CSV files.

REC_DIR,	which top level directory containing raw recordings, e.g., C:\data\nm07\recorded\
REDUCE_DIR,	which top level directory containing reduced data, e.g., C:\data\nm07\reduced\
CSV_DIR	e.g., C:\data\nm07\csv\

2. Create a spreadsheet of each station and channel recorded in the experiment.

   These data should be stored as columns in the spreadsheet.

   It is also useful to create a column containing the distance of each sensor from the source.

   Save this spreadsheet as: USER_STATION.CSV

3. Auto-generate a spreadsheet containing all the files to process. Run indexFiles.

   This creates GEN_FILE_INDEX.CSV, a spreadsheet of all the files, durations, # of channels, etc.

   The user must inspect this output.

4. Create a list of Yokogawa files. (This step is performed by the user.)

   Open GEN_FILE_INDEX.CSV, and eliminate all the files that are not Yokogawa blast recordings.

   Add column "irigb," which contains the IRIG-B channel number for each file.

   Save this file as USER_YOKO_FILES.CSV.

5. Determine the blast times.

   Run redStep1

   This creates GEN_BLAST_IRIGB.CSV, which is a list of all the IRIG-B times found.

   The user must inspect the file, add or remove IRIG-B times as necessary, and save as USER_BLAST_IRIGB.CSV.

6. Generate a spreadsheet containing the IRIG-B ranges recorded for all the files.

This is useful for finding bad IRIG-B timecodes and for speeding up the reduction process by ignoring files not containing the blast IRIG-B timecode.

Run `redStep2`.

This creates GEN_IRIGB_RANGE.CSV, which is a list of all IRIG-B ranges for each; it uses '.\cache' as a temp directory.

The user inspects the file and removes any entries with the IRIG-B signal 00-000-00-00-00, indicating "No signal found."

7. Reduce the data.

Run redStep3.

REDUCE_DIR is the target directory for reduction.

Each blast IRIG-B time will have a directory in this path.

The data corresponding to each station-channel will be a stored as a.blast and.meta file under the BLAST_IRIGB directory, e.g.,
\NM07\Reduced\07-223-22-40-1\A1-1.blast
\NM07\Reduced\07-223-22-40-1\A1-1.meta

Steps for calibration

1. User takes `GEN_FILE_INDEX.CSV` and creates a new CSV containing *only* calibration files.

   This file is saved as `USER_CAL_FILES.CSV`

2. Run `calStep1`

   This `preprocesses` the calibration waveforms to make identifying calibration information simpler.

3. Run `calStep2`.

   The user pro`vides the cached preprocess directory from reduction step 6 (above).`

   The file `GEN_CAL_INFO.CSV` is created, which contains all calibration information found.

4. The user takes the information in `GEN_CAL_INFO.CSV`, and identifies the `calibrations` to use.

   The user generates a new spreadsheet, `USER_CAL_RULES.CSV`, which may be from `GEN_CAL_INFO.CSV`, or entirely from scratch. `USER_CAL_RULES.CSV` is a listing of calibration rules to apply to the reduced waveforms. Each rule has an effective start and stop time, in IRIG-B, and it is inclusive.

   If a user chooses to create the calibration rules from scratch or modify `GEN_CAL_INFO.CSV`, the following are the column requirements:

station channel cal_scale cal_offset
cal_unit irigb_start irigb_stop

For example, a row in the calibration rules file may look like this:

`Y, 1, 120, 0, 'Pascals', 08-001-00-00-01, 08-002-00-00-01,`

which translates to: all "Y" stations, for Channel "1," between "08-001-00-00-01" and "08-002-00-00-01" inclusive, shall be calibrated with 120 mv/Pa, no offset, with the resulting waveform having the units of Pascals.

5. Run `validateCalRules`

This will check for overlapping effective times, as well as negative effective time durations. Errors will be reported to the MATLAB command line. The user should correct these errors. If errors occurred, the user should re-run `validateCalRules` again after corrections are made.

6. Run `calStep3` to apply the calibration rules to the reduced waveforms

Verification of the data reduction process

The final step checks all the reduced waveforms to ensure calibration information has been applied. This is done by running the `verify Reduction` function. At this point, it is also useful to make a list of the available channels per blast to identify stations that have stopped recording; this is done by running the `verifyReduction` function as well.

The output, `GEN_VERIFY.CSV`, will have a column for the IRIG-B time of the blast, and columns for all the station-channel pairs available. If calibration has been applied, an "`ok`" will populate the column. If the station was not recorded, a "`MISSING`" will be displayed. If calibration is missing, then "`uncal`" will be displayed.

For example, a row in the CSV file may read:

irigb	C-8	Y-1	A1-1
08-001-00-00-00	MISSING	ok	uncal

which means that, for blast "08-001-00-00-00," C-8 was not found; Y-1 was found and has been calibrated; and A1-1 was found, but not calibrated.

Once the calibration information is collected and verified, it is possible to separate the calibrated waveforms into separate files. This is accomplished by running `waveprofileFiles`, which can take days to complete, depending on the amount of data to be processed.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE (DD-MM-YYYY) 26-09-2012

2. REPORT TYPE Final

3. DATES COVERED (From - To)

4. TITLE AND SUBTITLE
Data Preparation Procedures for the ERDC High-Energy Large-Scale Blast Sound Propagation Experiment

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT

6. AUTHOR(S)
Daniel P. Valente, Lauren Ronsse, Roger D. Serwy, Jesse Barr, Keran Claffey, Michael J. White, and Michelle E. Swearingen

5d. PROJECT NUMBER 622720D048

5e. TASK NUMBER

5f. WORK UNIT NUMBER C43279

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
US Army Engineer Research and Development Center (ERDC)
Construction Engineering Research Laboratory (CERL)
PO Box 9005,
Champaign, IL 61826-9005

8. PERFORMING ORGANIZATION REPORT NUMBER
ERDC TR-12-12

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
US Army Public Health Command (USAPHC)
Building E1675, Room 210
Aberdeen Proving Ground, MD 21010-5403

10. SPONSOR/MONITOR'S ACRONYM(S) MCHB-IP-EON

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEME
Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES

14. ABSTRACT
This work generated acoustic data for nearly 28,000 blast events in a variety of weather conditions, and for associated meteorological conditions. Methods were developed to reduce and quality-check those data to create a dataset of over 18,000 curated blast waveforms, associated acoustic metrics, and corresponding meteorological data. This dataset provides a resource unique to the US Army Engineer Research and Development Center (ERDC) that may be used to assess the statistical variability of blasts noise as a function of distance, climate, and meteorology, and to investigate the definition of acoustic propagation classes for blast noise.

15. SUBJECT TERMS
atmospheric conditions, blast noise, military training, sound propagation

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			19b. TELEPHONE NUMBER (include area code)
Unclassified	Unclassified	Unclassified	SAR	32	