Blast noise classification with common sound level meter metrics

R. Cvengros
US Army Corps of Engineers, ERDC-CERL
Robert.Cvengros@usace.army.mil

D. Valente
US Army Corps of Engineers, ERDC-CERL
daniel.p.valente@usace.army.mil

E. Nykaza
US Army Corps of Engineers, ERDC-CERL
edward.t.nykaza@usace.army.mil

J. Vipperman
University of Pittsburgh
jsv@pitt.edu

December 12, 2011

Running Title: Blast noise classification with SLM metrics

Abstract

The study analyzes a common set of features that is measureable from a basic sound level meter (SLM), and attempts to understand the quality of mutual information carried in subsets of these features for the discrimination of military blast and non-blast classifications. Further, it seeks to develop intuition about how blast and non-blast feature vectors are distributed in high dimensional space by using the orthogonal centroid dimension reduction technique, which allows visualization in 2-dimensional space. Lastly, the study implements linear and Gaussian radial basis function (RBF) support vector machine (SVM) classifiers to determine the ability of this feature set to separate the two classes, as well as recursive feature elimination (SVM-RFE) to eliminate features that contain redundant information. The analysis is based on over 120,000 human classified signals in a combined dataset from ERDC-CERL and the University of Pittsburg. The accuracy of the linear and RBF SVM classifiers are listed for each of the experiments in the dataset, and the weights are given for the linear SVM classifier.

PACS numbers: 43.60.Bf, 43.60.Cg, 43.60.Np, 43.58.Gn

I INTRODUCTION

Monitoring the impact of noise sources is a fundamental necessity in environmental noise control engineering. Depending on the level of detail that is required for a given application, monitoring can involve recording an entire pressure time series, or can simply involve recording a set of metrics from a sound level meter (SLM). The advantages of using SLMs is that they are less expensive and require less data storage than recording equipment, in addition to requiring minimal downstream processing of recorded data. A down side of using SLMs for noise assessment, however, is that this type of monitoring typically requires human listeners to verify or annotate the signals present during data capture, specifically to ensure that the levels being monitored are associated with the source of interest. This is not a significant challenge for short data collections in a limited number of monitoring locations, however, it quickly becomes an issue when the assessment duration or number of monitoring locations increases. As unattended monitoring networks become increasingly attractive to the noise control engineer, one may ask whether a decision criteria could be developed based on the output of an SLM to decide whether a source of interest was present during a given assessment period. If so, one could automate the task of assessment to a large degree, allowing for real-time, continuous noise monitoring which requires little human intervention. This article discusses precisely this task, the classification of waveforms given only metrics recordable on a typical SLM.

A specific application of noise monitoring for which this problem is relevant is in monitoring blast noise produced on military installations. Blast noise is typically associated with the sound produced by large weapons and artillery, and is characterized by high amplitude, low frequency energy that can retain significant amplitudes at long distances. Military installations typically monitor these levels in the surrounding community to inform their testing and training decisions and to limit disturbance of the residents living nearby. For monitors running continuously on an installation's perimeter or in surrounding communities, a high volume of data are generated. Typically, the monitor is triggered by an event exceeding a threshold level; therefore, these data may contain a large number of non-blast recordings, most notably due to wind noise, nearby vehicles, physical manipulation of the microphone or the mounting structure, or other loud amplitude sounds. As a result, it is useful to have an algorithm which can sift through the data and reliably classify incoming signals, thus aiding in proper determination of blast noise levels, assessment of the impact of specific events, and the future analysis of recorded datasets.

Several authors have previously approached the problem of blast noise classification. Bucci and Vipperman (2007) used an artificial neural network (ANN) to classify blast noise waveforms obtained from noise monitors. Their classifier was based on two time domain features, the crest factor (peak to RMS value) and the kurtosis (standardized fourth moment of the waveform distribution), along with two frequency domain features, the slope of the log power spectral density (PSD) within the 0-100Hz bandwidth of interest, and the residual error when a line of this slope is

fit to the PSD. While the time domain metrics are generally correlated with impulsive signals, the frequency domain metrics were used to appropriately classify non-blast noise sources such as wind and aircraft noise. These metrics were calculated for a training set of 1000 signals (330 military impulse, 670 non-impulse). They trained and tested the radial basis function (RBF), self-organizing map (SOM), and multilayer perceptron (MLP) ANN classifier structures, as well as a linear least squares classifier for comparison. They found that, while the non-linear ANNs were able to achieve 98.2-100% accuracy with all four metrics, the least squares classifier performed with 94.1% accuracy.

Rather than attempt to develop an optimal classifier from designed metrics, the current study analyzes a more common set of features measureable from a basic SLM. We attempt to understand the quality of information carried in subsets of these features, and how these contribute to the discrimination of blast and non-blast signals. Further, we investigate the distribution of blast and non-blast feature vectors in high dimensional space by using the Orthogonal Centroid algorithm for dimension reduction (Kim *et al.*, 2005), allowing data to be visualized in two dimensions for ease of visualization and interpretation. Lastly, this study implements linear and Gaussian RBF support vector machine (SVM) classifiers (Cortes and Vapnik, 1995) using the SLM feature set, and eliminates features containing redundant information using recursive feature elimination (SVM-RFE) (Guyon *et al.*, 2002). The analysis is based on over 120,000 human classified signals in a combined dataset of blast and non-blast waveforms from different locations, different conditions, on different monitoring equipment. The accuracy of the linear and RBF SVM classifiers are examined for each of the experiments in the dataset, and the weights describing the importance of each feature in the discrimination are given for the linear SVM classifier.

II   DATASETS

This study is comprised of data from six different locations in seven separate data collections. The combined data set contains 60,916 blast signals and 59,574 non-blast signals, for a total of $M = 120,490$ total human classified signals.

A   ERDC Long Range Propagation Experiments (LRPE)

This dataset was obtained in a set of controlled experiments with the expressed purpose of understanding the variability in blast noise levels as a function of distance from the source (Ronsse *et al.*, 2011). The experiments took place at two locations in the United States at two times of year. The first experimental location was in an arid desert climate with minimal terrain features and vegetation. The other test location was in a temperate climate, which was forested and moderately hilly. The blast sources for these experiments were 1.25lb blocks of C4, detonated at twenty minute intervals over 6-10 hour test periods. These test periods were scattered at different times of day over the 2-3 week duration of each experiment so as to sample a variety of meterological conditions at different times of day. At each location, tests were carried out in the summer and the winter in order to measure seasonal variation in received level.

3

Blast waveforms were measured at a variety of distances out to 16km in three different directions from the source. Blast waveforms were acquired with 1/2-inch diameter Brüel and Kjær Type 4921 microphones mounted 1.5m from the ground, and recorded on Rion DA-20 4-channel digital data recorders, which continuously ran over the course of each test period. Using windows based on the expected time of arrival for each blast at each monitoring location, over 28,000 waveforms were extracted and 19,720 were valid and classifed by human listeners as containing a blast or not. The LRPE dataset includes 18,303 human classified blast signals and 1,417 non-blast signals. Absence of blasts from the waveforms is due to meteorlogical conditions or high ambient noise environments. A more detailed description of these experiments and the data cleaning procedures can be found elsewere (Ronsse *et al.*, 2011).

The data analyzed in this article include all waveforms acquired at distances of 1km, 4km, 8km, 12km, and 16km in each experiment. The non-blast signals contain a variety of other noise sources: ground vehicles, aircraft, small arms fire, wildlife, speech, and most notably, wind.

B   East South Central (SERDP–CERL)

This dataset was collected in support of a large human response to blast noise project (Valente *et al.*, 2011). Data collection took place in one location in the east south central region of the United States over the course of 12 months. The location's geography features a temperate forest climate and moderately hilly terrain. The test location was a military installation where small and large arm training exercises occurred daily.

The recorders were located both on the installation as a set of five-microphone arrays, and off the installation outside of local residents' homes as single microphone posts. The monitors were designed to record environmental noise for 5 seconds given a stimulus that surpassed the 100 dB SPL trigger level, and included a 0.5 second pre-trigger time. The on-installation trigger microphones were part 377A13–NR (PCB Piezotronics) 1/2-inch diameter micro-phones, while all other microphones were part 377B11–E (PCB Piezotronics) 1/2-inch microphones. On-installation array microphones were placed cylindrically with quasi-random elevation below 2.5 m and quasi-random angle. Off-installation single microphones were placed at 2–3 m elevation. All waveforms were recorded using HBM Liberty data acquisition systems.

Blast sources were located on a closed set of designated firing points over a ~ 450 km^2 area. Receivers were approximately $0.5 - 20$ km from the source, depending on the source and receiver pair. Approximately 3 million signals were recorded during the course of the data collection, however, only 4,782 valid signals were classified by human listeners. The SERDP–CERL dataset includes 4,433 human classified blast signals and 349 non-blast signals. Non-blast signals are composed other noise sources: wind, thunder, wildlife, human, electronic, ground vehicles, aircraft, and small arms fire.

C Aberdeen Proving Ground (APG1 & APG2)

This dataset was obtained in support of a complaint risk study (Nykaza *et al.*, 2008) conducted by the Construction Engineering Research Laboratory (CERL) between 2006 and 2008. Ten unattended Nor121 data acquisition systems where located near major firing points to help researchers determine the source locations. The distance between source(s) and the recorders typically ranged between 500 meters to 5 kilometers. Given that the experiment was conducted over a two year time period, the waveforms were recorded for a variety of blast producing noise sources (e.g., $\geq$ 50 mm guns and various large explosives). Similar to the procedure used in the ERDC LRPE, research assistants listened to the recorded waveforms and notated whether the recorded event was a blast. In addition to notating blast events, non-blast events or additional distinguishable non-blast sounds that occurred at the same time of blast events were notated. Non-blast signals are composed other noise sources: wind, thunder, wildlife, human, electronic, ground vehicles, aircraft, and small arms fire.

This dataset represents two separate data collections, APG1 and APG2, where the SLM monitor locations were the same for both experiments. The APG1 dataset contains 5,407 human classified blasts and 19,120 non-blasts . The APG2 dataset contains 22,081 human classified blasts and 21,985 non-blasts.

D New York (Fort Drum)

The Fort Drum dataset includes 6,211 human classified blast signals and 9,545 non-blast signals.

E SERDP-PITT

The University of Pittsburgh (Pitt) developed a high fidelity library of recorded noise sources found around military installations. Waveforms were identified and recorded in a log at the time of the recording. Pitt worked with the Range Control office at each base to identify locations and sources of noise. The library contains a total of 2,471 waveforms, including 740 impulse (blast) and 1,731 non-impulse signals (aircraft, wind, vehicle noise, etc.).

This measurement setup was based upon a Larson Davis NMS-011 Environmental Noise Monitoring System. The primary microphone was replaced with a Bruel & Kjaer (B&K) 4193 infrasonic microphone, which has a bandwidth of 70 mHz to 20 kHz. The primary microphone was connected via cable to a Larson Davis (LD) 824 Sound Level Meter (SLM). The LD 824 served as a field-portable microphone power supply that also logs Leq and Lpk value at one second intervals (for subsequent data validation steps). Since most of the energy of the sources to be measured lies in the very low-frequency range (0-100 Hz) the weighting of the input spectrum is set to flat (linear) weighting. The AC output of the LD 824 SLM was connected to a National Instruments (NI) DAQCard-6036E data acquisition card through a NI BNC-2110 input/output board. When weather permitted, a second channel was recorded using a LD 2540 microphone with a B&K 5935 power supply. The DAQCard was installed into a Dell Latitude laptop with a

5

Pentium IV processor. A Virtual Instrument (VI) was created in Labview 7.1 to capture waveform data. The VI enables an automatic triggering/pre-triggering mode, where data are recorded when the signal exceeds a specified threshold, (used to automatically record impulse events that exceed a certain Lpk value). Although the threshold can be set to record at or above any desired Lpk level, it was typically adjusted to just above ambient noise levels (typically 80-95 dB) in order to record as much data as possible. In the data collection, a 0.1-0.25 second pre-trigger was coupled with an additional 2 seconds of recorded data for each record. By pre-triggering, the entire discrete event was able to be recorded. A manual (continuous) triggering mode was also possible, which was used to record longer or continuous events such as wind, aircraft noise, traffic, and engine noise. During multiple successive trigger events, the automatic mode also triggered nearly continuously. The data were sampled at 10 kHz.

Eleven (11) measurement trips were made to seven (7) locations. Trips were conducted across all seasons and weather conditions and also provided a wide variety in topology and vegetation. Table I summarizes the measurement locations and the types of noise measured there.

Where present, measurements were conducted in the vicinity of permanent monitoring stations, in order to get accurate representations of the signal that the monitors try to classify. In addition, the barometric pressure, date, time, and weather conditions were recorded in the test log.

F   North Carolina (MCBCL-PITT)

Data measurements made a Marine Corps Base Camp Lejeune (MCBCL) were performed by two prototype noise monitoring systems called the Bearing Amplitude Measurement and Analysis System (BAMAS) [Applied Physical Sciences, 2009]. The electronics are composed of four PC/104 modules, PC/104 ISA extension card, and rechargeable Li-Ion battery. The data acquisition board is capable of sampling 16 channels, with 16 bit resolution, at 200kHz. Four array microphones and one high-quality ACO PACIFIC, INC 7052S condenser microphone with an ACO PACIFIC 4052 preamp (4-22k Hz) simultaneously recorded events. It was the data from the ACO microphone that was used for this study. The microphone is capable of measuring levels up to 140 dB. All channels were sampled at 5 kHz and were filtered by a 6th order low-pass Butterworth filter with a corner frequency of 2.5 kHz.

One prototype was installed near a distant existing noise monitor and receives few blasts. The other was installed near an observation point. The observation point is close to the impact zone (as little as 0.5 km), but is distant from the firing positions (10 km). This location makes the monitor susceptible to operational noise (especially vehicles and electronic noise) and small arms fire that more distant noise monitors wont receive. While the data arent as pure as the SERDP-PITT library, they represent realistic measurements under typical conditions.

A person listened to each recorded waveform, plotted it along with its spectrum, and looked at detected output from the BAMAS array (could better see acoustic events, since incoherent (wind) noise was canceled by the array

processing). Of the 9,168 useful waveforms, 3,741 were blasts and the remaining 5,427 were not. The data were processed to assess and compare operation of the algorithms.

III METHODS

A Metric Calculations

The metrics analyzed in this manuscript include those defined in the ANSI S1.4-1983 (Acoustical Society of America, 1983) specification for Sound Level Meters, along with frequency weighting filters as specified in ANSI S1.42-2001 (Acoustical Society of America, 2001). The complete set of metrics are: Apeak, Cpeak, Zpeak, LAMAX fast, LCMAX fast, LZMAX fast, LAMAX slow, LCMAX slow, LZMAX slow, ASEL, CSEL, ZSEL, ALEQ, CLEQ, and ZLEQ. The time and frequency weighting filter definitions used in this study, as well as definitions of sound exposure, sound exposure level, and equivalent-continuous sound level are shown in Table II.

The frequency-weighting filters are implemented in MATLAB using frequency domain filtering techniques.

$$w[n] = iFFT\{H(f) \cdot FFT\{x[n]\}\} \tag{1}$$

The input waveform was first zero padded to $2^{ceil(\log_2(length(x)))+1}$ to avoid circular convolution artifacts. The complex frequency response $H(f)$ was determined using MATLAB's `freqs` command, with the analog poles and zeros given in Table II. The resultant signal was then truncated to the original signal length with negligible effect on the waveform.

The time-weighting filters are implemented using the infinite impulse resonse (IIR) filter

$$H(z) = \frac{B(z)}{A(z)} = \left.\frac{\alpha}{1-(1-\alpha)z^{-1}}\right|_{\alpha=\frac{1}{\tau \cdot Fs}} \tag{2}$$

$$y[n] = \alpha \cdot x^2[n] + (1-\alpha) \cdot y[n-1], y[0] = x[0] \tag{3}$$

$$w[n] = 10\log_{10}\left(\frac{y[n]}{p_0^2}\right), \tag{4}$$

where $x[n]$ is the input signal and $w[n]$ is the time-weighted output signal. The impulse response of this IIR filter was verified to be equivalent to the explicit defininition from Table II. This method of time weighting is preferred over the explicit form due to the great increase in computational efficiency and for its accuracy to the specification that comes as a result of not truncating the impulse response of the filter.

The set of metrics described above were calculated for each waveform in the datasets described in Sec. II. Each signal is therefore represented by a feature vector $\mathbf{x} \in \mathbb{R}^{15}$ and a human classified value of 'blast' or 'non-blast'.

7

B Support Vector Machine classifier

The human classified datasets described in Sec. II provide M human classified waveforms, each represented by a feature vector ξ_i where $i = 1, \ldots, M$. Each waveform is marked by human listeners with a label $y_i \in \{-1, +1\}$, where $y_i = +1$ corresponds to a 'blast' and $y_i = -1$ corresponds to a 'non-blast'. The set of blast points and non-blast points form distributions in $\mathbb{R}^{15}$ space. It is the goal of this manuscript to develop a decision rule which optimally separates these two distributions. Support Vector Machines (SVMs) are one way to specify this boundary, and can be formulated for either linear boundaries or non-linear boundaries by using non-linear kernels (Cortes and Vapnik, 1995).

SVMs attempt to optimally separate two distributions with boundary $\mathbf{w}^T\mathbf{x} + b = 0$ by maximally extending margins $\mathbf{w}^T\mathbf{x} + b = \pm 1$ on either side of the boundary hyperplane such that all blast points satisfy $\mathbf{w}^T\mathbf{x} + b \geq 1$ and all non-blast points satisfy $\mathbf{w}^T\mathbf{x} + b \leq -1$. The problem can be stated as an optimization problem that maximizes the distance between margins $\gamma = \frac{2}{||\mathbf{w}||}$ subject to the inequality constraints.

The classification rule used to determine the class of test point $\mathbf{z}$ is ultimately given by,

$$f(\mathbf{z}) = \mathbf{w}^T\mathbf{z} + b \tag{5}$$

$$f(\mathbf{z}) = \sum_{i=1}^{M} \alpha_i^* y_i \langle \mathbf{x}_i, \mathbf{z} \rangle + b \tag{6}$$

$$f(\mathbf{z}) \begin{cases} > 0 & \mathbf{z} \text{ is classified as a blast} \\ < 0 & \mathbf{z} \text{ is classified as a non-blast} \end{cases} \tag{7}$$

where α_i^* is a constant obtained from the optimization procedure and is only non-zero for support vectors, and $\langle \cdot \rangle$ denotes the inner product. A finite cost C for points crossing the boundary can be enforced during optimation, which manifests as a limit on the value for α_i^*.

The SVM formulation can be modified to extend SVMs as non-linear classifiers. The inner product in Equation 6 can be considered a "similarity" function that is a high value when vectors $\mathbf{x}_i$ and $\mathbf{z}$ are similar and approaches 0 when the two vectors are dissimilar. The inner product can be replaced by other kernel functions $K(\mathbf{x}_i, \mathbf{z}) = \phi(\mathbf{x}_i)\phi(\mathbf{z})$ that convey similarity. Two commonly used kernel functions are described in Table III. Functions $\phi(\cdot)$ create separation between two distributions by mapping each point to a higher dimensional space. For, example the RBF kernel maps the data to an infinite dimensional space. As a result, it is impossible to represent mapped vectors $\phi(\mathbf{x}_i), \phi(\mathbf{z})$ in software, however it is quite efficient to represent the combination $K(\mathbf{x}_i, \mathbf{z})$.

In this analysis, an SVM classifier with cost $C = 10$ is implemented using the linear and RBF ($\gamma = 0.1$) kernels. The performance of these classifiers is evaluated in terms of the jackknife cross validation accuracy as discussed in the next section. All data processing routines were written in Python using the PyML module (`http:`

//pyml.sourceforge.net/), which is based on the libsvm (http://www.csie.ntu.edu.tw/~cjlin/libsvm/) library.

C   Data Jackknife Procedure

Due to the large size of the dataset, a jackknifing procedure was developed to mitigate classifier training time. The SVM was trained from a subset of feature vectors that were randomly sampled from the dataset and tested on the remaining feature vectors to assess classifier accuracy. The classifier performance was evaluated by measuring the mean and standard deviation of the accuracy over 50 random samplings of the data to ensure the stability of that statistic.

D   Recursive Feature Elimination

Features which contribute little to the ability for the SVM to discriminate between classes can be eliminated in a procedure termed Recursive Feature Elimination (SVM-RFE) (Guyon *et al.*, 2002). This procdure can also provide a means for feature ranking and selection. The algorithm recursively iterates through the following steps:

1. Train the linear SVM classifier to get optimal feature weights, w_i

2. Compute the ranking criterion w_i^2 for each metric

3. Eliminate the feature with the smallest ranking criterion

This process was used to rank the features calculated in Sec. III-A with respect to their ability to distinguish blasts from non-blasts.

SVM-RFE feature ranking was executed 40 times on jackknifed samplings of 5000 training vectors. The probability of each feature lying within a given ranking $P(R)$ was calculated and the expected value of the rank $E[P(R)]$ was then used to determine a master ranking for all the features.

E   Orthogonal Centroid algorithm for Dimension Reduction

In an effort to better understand which metrics best represent the impulsivity of the blast as well as how the blast and non-blast points are distributed in $\mathbb{R}^{15}$ space, dimension reduction was performed on the data. The aim was to represent the $\mathbf{x}_i \in \mathbb{R}^{15}$ vectors as $\psi(\mathbf{x}_i) \in \mathbb{R}^2$ vectors so that the distributions could be easily visualized.

For dimensionality reduction, the orthogonal centroid algorithm (Kim *et al.*, 2005) was chosen for its propensity to retain clustering information within the reduction to $\mathbb{R}^2$. The algorithm requires the centroid (mean) for each the

9

distributions, $\mathbf{c}_b$ and $\mathbf{c}_{nb}$, where $\mathbf{c}_b$, $\mathbf{c}_{nb}$ are 15×1 dimensional vectors. Then

$$C_{[15 \times 2]} \quad = \quad \begin{bmatrix} c_{b[15 \times 1]} & c_{nb[15 \times 1]} \end{bmatrix} \tag{8}$$

$$C_{[15 \times 2]} \quad = \quad Q_{[15 \times 2]} R_{[2 \times 2]} \tag{9}$$

$$\hat{\mathbf{x}}_{[2 \times 1]} \quad = \quad Q^T_{[2 \times 15]} \mathbf{x}_{[15 \times 1]} \tag{10}$$

Bracket notation is used in each variable's subscript to denote the size of the vector/matrix for ease of interpretation. Equation 9 represents the rectangular QR matrix decomposition into orthonormal basis Q and upper triangular matrix R. Further, the boundary between the blast and non-blast clusters in $\mathbb{R}^{15}$ space defined by the linear SVM described in Sec. B can also be represented in this reduced $\mathbb{R}^2$ space

$$\hat{\mathbf{w}}_{[2 \times 1]} \quad = \quad Q^T_{[2 \times 15]} \mathbf{w}_{[15 \times 1]} \tag{11}$$

$$\hat{\mathbf{w}}^T_{[1 \times 2]} \hat{\mathbf{x}}_{[2 \times 1]} + b \quad = \quad 0 \tag{12}$$

Inherent in the dimesion reduction is a loss of information and an associated classifier accuracy loss, however, intuition is gained as a result of being able to visualize the data in $\mathbb{R}^2$ space.

IV   RESULTS

A   Support Vector Machine classifier

The size of the jackknife sampling was varied from $10 - 10,000$ training vectors for the linear kernel and from $10 - 60,245$ for the RBF kernel, and classifiers based on each sample size were evaluated from 50 random samplings. The range of jackknife sizes for the linear kernel is limited due to the extended training time required as compared to the RBF kernel.

The mean and standard deviation of the accuracy was tracked as the sample size varied; these curves are shown in Figure 1. These statistics were found to saturate near an accurancy of $91.2\% \pm STD$ for the linear classifier at a sample size of 1000. As a result, all analyses involving a linear kernel in this study are trained on a random sample size of 1000 or greater. The RBF classifier accuracy did not saturate at any jackknife sample size due to the flexibility of the non-linear boundary. The accuracy continued to increase as size of the training set increased, and achieved a mean accuracy of 95.3% when the classifier was trained with half the dataset and tested on the other half.

The SVM classifier accuracies for the linear and RBF kernels, when trained with 5000 feature vectors randomly sampled from the entire dataset, are shown in Table IV. Over the combined dataset, the linear classifier performed
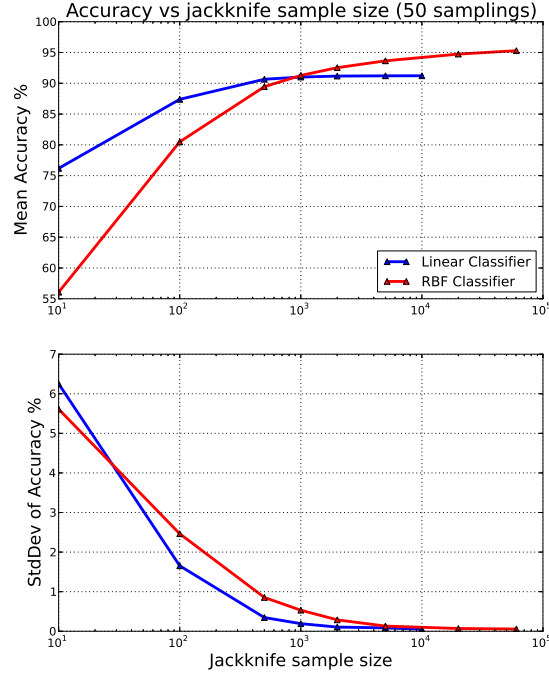
Figure 1: The mean and standard deviation of accuracies of several jackknife sample sizes from 50 samplings. Note that the the mean and standard deviation saturate above sample sizes of 1000 for the linear classifier at 91.2% accuracy. The mean and standard deviation do not saturate for the RBF classifier, and achieve 95.3% accuracy when the jackknife sample size is half the size of the entire dataset.

with 91.3% accuracy and the RBF classifier performed with 93.8%. Using this classifer, blasts from all experiments except for Fort Drum were classified with greater than 90% accuracy with the linear classifier, while the Fort Drum classifier accuracy was 82.8%. APG1, APG2, LRPE, MCBCL-PITT, SERDP-PITT, and SERDP–CERL had greater than 90% accuracy with the RBF classifier, while the Fort Drum experiment was classified with 86.5% accuracy.

B   Orthogonal Centroid algorithm

The distributions of blast and non-blast feature vectors for each experiment are shown in Figure 2 in the reduced-dimension coordinate space $(\hat{x}_1, \hat{x}_2)$, as described in Sec. E. The linear SVM boundary $\mathbf{w}$ is trained on the full dimensional data and projected to $\mathbb{R}^2$ using the same technique used to reduce the dimension of each point.

For experiments APG1, APG2, SERDP-PITT, and SERDP–CERL the blast and non-blast distributions are fairly separable. The LRPE experiment shows a region where blasts and non-blasts overlap completely, as well as a region where there are blasts but no non-blasts. The Fort Drum and MCBCL-PITT experiments show a high degree of overlap between the two distributions.

The combined dataset shows two distinct sub-clusters within the non-blast category that appears to be experiment
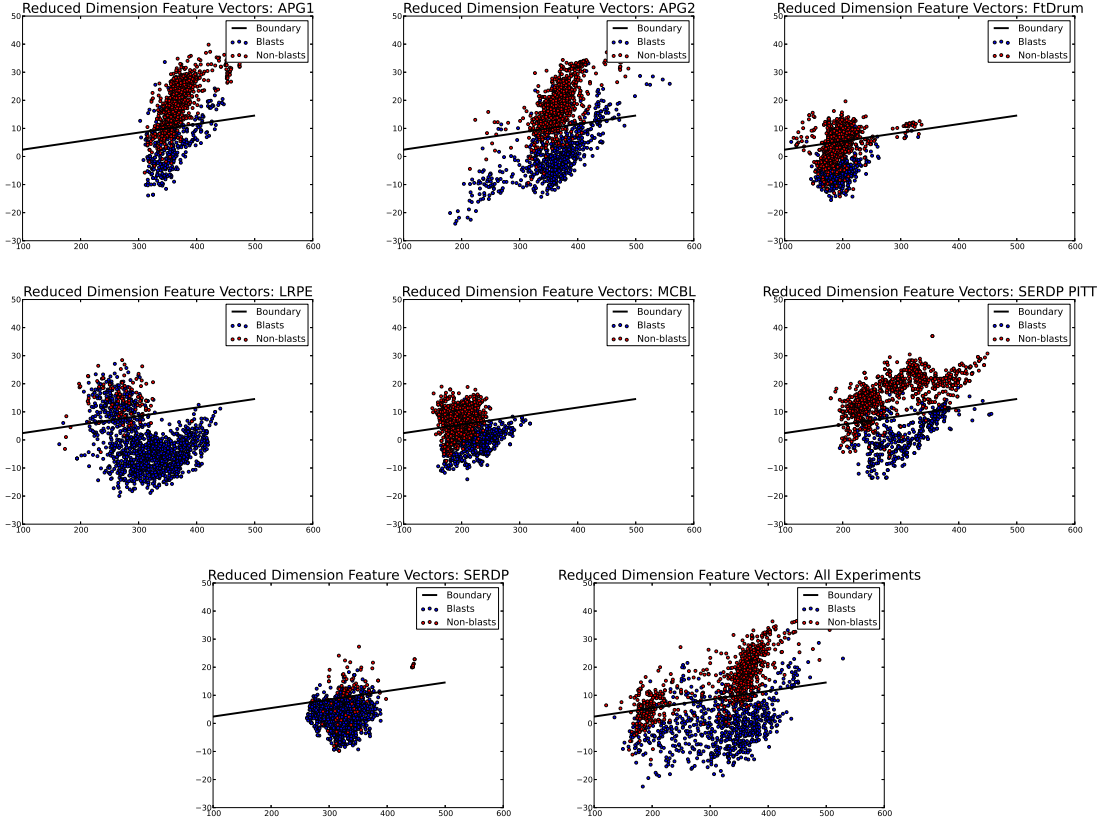
11

Figure 2: The sampled and reduced dimension distributions in $(\hat{x}_1, \hat{x}_2)$ coordinate space. of blasts and non-blasts based on the orthogonal centroid algorithm for each data set. The reduced dimension linear SVM boundary is drawn for reference.

dependent; the right non-blast cluster from the combined dataset appears to be mostly composed of the non-blasts from APG1, APG2, and SERDP–CERL experiments, while the left non-blast cluster appears to be largely combosed of the Fort Drum, LRPE, MCBCL-PITT, and SERDP-PITT experiments. The two non-blast sub-clusters are seperated by a space that is not dominated by either category. The combined data from all experiments shows the two distributions may not be efficiently separable using a linear boundary; a large cluster of non-blast point lie on the wrong side of the boundary. The non-linear RBF boundary is capable of better isolating the two non-blast sub-clusters from the blast distribution, however, this method risks overfitting the data by incorrectly classifying the space in between the two non-blast sub-clusters.

C Feature Ranking and Selection

The features ranked by the SVM-RFE algorithm anre listed in order of importance in Table V, along with the expected rank value given the rank placement probabilities from the 40 independent jackknife rankings (see Sec. D). The feature ranking probabilities are shown for each feature in Figure 3. The feature set is ordered by ascending expected
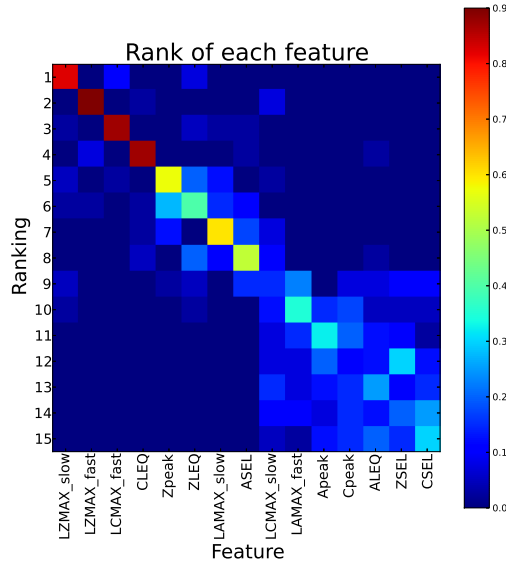
Figure 3: The probability that a feature is ranked at a given rank. Probabilities are calculated based on 40 SVM-RFE rankings of jackknife samplings of size 5000. Features are ordered by ascending expected rank values from left to right.

rank values from left to right. The expected rank values seem to be clustered in groups; a high value range from ranks 1–3, a medium value range from ranks 4–8, and a low value range from ranks 9–15.

This ranking was tested by iteratively eliminating the least signifcant features and evaluating the resultant linear and RBF classifier accuracies over the combined-experiment dataset. The result of this analysis is shown in Figure 4. The mean squared difference between the RBF and linear SVM classifiers was 2.5%. In general the RBF SVM classifier is more accurate than the linear SVM classifier, however, the two cases converge as many features are eliminated. It is important to note that the relatively low mean-squared difference between the linear and RBF curves indicates that the optimal SVM boundary is approximately linear. Therefore, the intuition gained by looking at the weight vector $\mathbf{w}$ of the linear classifier is justified.

A reasonable simplification of this classification problem would be to implement the SVM classifier using only a few top-ranked features. This would trade classifier performance for the convenience of a small feature set. In some applications, this trade-off may be desirable. Table VI shows the feature weights $\mathbf{w}$ for each of the top features, the linear bias b, and the resultant blast/no-blast classifier accuracy, true positive rate (TPR), and false positive rate (FPR). Figure 4 shows that it is not useful to include more than eight features in the SVM, since doing so will not result in better classifier performance. The weights in Table VI can be used to test a new signal feature vector $\mathbf{z}$ using equations 5 and 7.

We can further analyze the feature weights from Table VI to learn which features "push" toward the blast or non-

blast classifications based on their sign and magnitude. From Table VI we can see that, with the exception of Zpeak, the negatively weighted features (those pushing towards "non-blast") are all indicators of background level. The postively weighted features (those pushing toward "blast") indicate the max signal level using a fast time weighting. Thus, signals that are most likely to be classified as blasts have, in general, high max levels given a fast time weighting and a low measured background level.
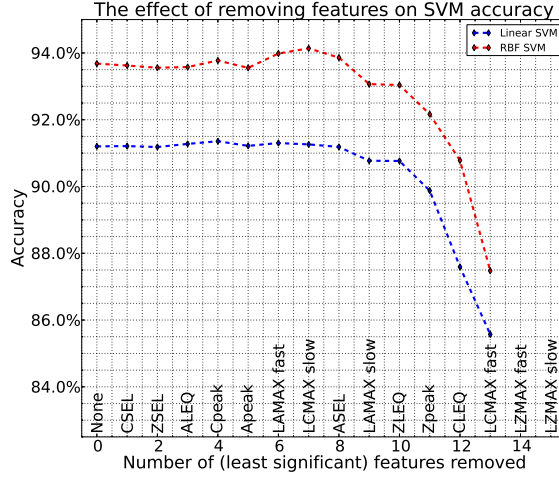

Figure 4: The effect of removing least significant features from the feature set on the linear and RBF SVM accuracy. The mean squared difference between the RBF and linear SVM classifiers was 2.5%. The feature removed at each horizontal step is listed above is located above the abscissa.

Table VI also shows that an 85.6% accuracy classifier can be created using just two metrics that have approximately equal weights

$$f(\mathbf{z}) = LZMAX\_fast(\mathbf{z}) - LZMAX\_slow(\mathbf{z}) \begin{cases} > 5.40 & \mathbf{z} \text{ is classified as a blast} \\ < 5.40 & \mathbf{z} \text{ is classified as a non-blast} \end{cases} \tag{13}$$

which has been normalized to unit feature weightings. This classifier can be valued for its simplicity, however, it is less effective than a classifier based on a larger feature set.

V   CONCLUSION

The datasets described in this article were obtained using vastly different equipment in vastly different settings. The compiled dataset of over 120,000 records provides an impressive compendium with which to analyze, in detail, useful features for discriminating blast signals from non-blast signals. As is common to many applications in machine learning, the particular features that provide optimal discrimination are not trivial to recognize. However, by restricting

14

these features to those obtainable from simple SLM recordings and reducing the dimensionality of the resultant feature vectors, we were able to gain insight into how blasts and non-blasts are distributed in this high-dimensional feature space, as well as how these distributions vary between datasets (Figure 2). Although the distributions contain some differences per experiment, the cumulative dataset contained trends that allowed a linear boundary created from the entire dataset to be reasonably successful for each experiment when considered on its own. For at least one dataset, the accuracy of the linear classifier exceeded 95%. A different set of features would undoubtedly increase classifier performace, for instance, adding in the kurtosis and crest factor will certainly improve accuracy. However, more complex features require complete pressure time series and computational power that is not always available in a given circumstance.

It is interesting to note that many of the metrics chosen in this study contain redundant information; they are only slightly different transformations of the original waveforms—some of which only amount to a marginally different filtering in the frequency domain (e.g., C and Z filters), and some which even show approximately linear correlations. Nevertheless, the slight difference in information provided by each of these transformations serve to accentuate differences between blasts and non-blasts enough so that removal of the apparently redundant features degrades classifer performance. This is illustrated in Table VI and Figure 4. Furthermore, the SVM-RFE algorithm quantifies our general intuition of how one would choose features to discriminate blast signals from non-blasts. Metrics which provide information about both the maximum level of a signal and its background level are useful discriminators of blast noise, and frequency weightings that cause the least amount of distortion on the signal are to be preferred. The linear classifier weights demonstrate the importance of each metric as it contributes to a blast or non-blast classification.

The non-linear boundry created by the Gaussian RBF kernel for the SVM classifier is, however, capable of achieving a higher accuracy than the linear classifier, especially when the size of the training set is increased. Theloss of interpretive value and simplicity, however, is significant. A field noise monitor could easily implement the RBF classifier and acheive an accuracy greater than 95%, but, again, the availability of computational power is limited in certain applications. Where "on the go" type classifiers are desired, the linear weights provide a quick and simple alternative with little degradation of results. In fact, if one is willing to accept a classifier with 85% accuracy, Equation 13 provides a simple linear classifier for determining the presense of a blast from only the difference between the fast- and slow-weighted maximum levels. Both linear and nonlinear methods offer a concrete solution to blast noise classification using sound level meters, but as is the case in any problem in engineering, implementation is at the discretion of the engineer.

References

Acoustical Society of America (**1983**). *American National Standard Specification for Sound Level Meters*, New York, New York, ANSI S1.4-1983.

Acoustical Society of America (**2001**). *American National Standard Design Response of Weighting Networks for Acoustical Measurements*, New York, New York, ANSI S1.42-2001.

American National Standards Institute (**1993**). *Methods for measurement of impulse noise*, New York, New York, ANSI S12.7-1986 (R1993).

American National Standards Institute (**1996**). *Impulse Sound Propagation for Environmental Noise Assessment*, New York, New York, ANSI S12.17-1996.

(APS), A. P. S. (**2009**). "SI-1427 Final Report: Impulse Noise Bearing and Amplitude Measurement and Analsis System", .

Benson, J. (**1996**). "A real-time blast noise detection and wind noise rejection system", Noise Control Eng. J. **44**, 307–314.

Bucci, B. A. and Vipperman, J. S. (**2007**). "Performance of artificial neural network-based classifiers to identify military impulse noise", J. Acoust. Soc. Am. **122**, 1602–1610.

Bucci, B. A. and Vipperman, J. S. (**2008**). "An Investigation of the Characteristics of a Bayesian Military Impulse Noise Classifier", in *NCAD2008* (Dearborn, MI).

Cortes, C. and Vapnik, V. (**1995**). "Support-vector networks", Machine Learning **20**, 273–297.

Guyon, I., Weston, J., Barnhill, S., and Vapnik, V. (**2002**). "Gene selection for cancer classification using support vector machines", Machine Learning **46**, 389–422.

Kim, H., Howland, P., and Park, H. (**2005**). "Dimension Reduction in Test Classification with Support Vector Machines", J. of Mach. Learn. Res. **6**, 37–53, centroid approach to dimensionality reduction.

Luz, G. A. (**1980**). *Suggested procedures for recording noise complaints at army installations*, United States Army Environmental Hygiene Agency, Aberdeen Proving Ground, MD 21010-5422, HSE-08/WP Technical Guide.

Nykaza, E., Pater, L., and Luz, G. (**2008**). "Improved procedure for correlating blast noise events with complaint logs at u.s.army installations", Noise Control Eng. J. **56**, 451–459.

371   Ronsse, L., Valente, D., Swearingen, M., White, M., Pater, L., MacAllister, B., Mifflin, J., Wilson, D., and Albert,

372   D. and, J. B. (**2011**). "Blast noise characteristics as a function of distance and climate", J. Acoust. Soc. Am. .

373   Russell, W. A., Luz, G. A., and et. al. (**2001**). *Environmental Noise Management: An Orientation Handbook for Army*

374   *Facilities*, U.S. Army Center for Health Promotion and Preventive Medicine (USACHPPM).

375   Schomer, P., Bandy, M., Lamb, L., and Van Slooten, H. (**2000**). "Using fuzzy logic to validate blast noise monitor

376   data", Noise Control Eng. J. **48**, 193–205.

377   Schomer, P. and et. al. (**1994**). "Human and Community Response to Military Sounds: Results from Field-Laboratory

378   Tests of Small-arms, tracked-vehicle, and Blast Sounds", Noise Control Eng. J. **42**, 71–84.

379   Valente, D., Nykaza, E., Hodgdon, K., Gaugler, T., Krecker, P., MacAllister, B., and Luz, G. (**2011**). "Assessing the

380   impact of blast noise on communities near U.S. Army installations", in *11th International Congress on Noise as a*

381   *Public Health Problem (ICBEN)*.

Table I: List of measurement sites visited for SERDP-PITT data collections.

Location	Number of Visits	Types of Noise
MCBCL, NC	3	Tank, artillery, vehicle, mortars, wind noise, dummy bombs, demolition charges, grenades, Bradley fire, aircraft
Fort Indiantown Gap, PA	2	Mortars, Howitzer, dummy bombs, canon strafe, vehicle strikers
Fort Benning, GA	2	Tanks, demolition, Bradley fire, mortars, vehicle, aircraft, wind
Fort Hood, TX	1	Tanks, Palladin, Bradley, aircraft, mortars, artillery, vehicle, wind
Fort AP Hill, VA	1	Demolition, grenades, mines, mortars, artillery, aircraft, wind
Fort Riley, KS	1	Tanks, Bradley, demolition, aircraft, strong winds
Fort Carson, CO	1	Tanks, Bradley, demolition, mortars, aircraft, winds

Table II: Method definition for time weighting, frequency weighting, SEL, and LEQ.

Method	Definition
Time weighting	$L_\tau(t) = 10 \log_{10}\left(\frac{1}{\tau}\int_{t_s}^{t} \frac{p^2(\xi)}{p_0^2} e^{\frac{-(t-\xi)}{\tau}} d\xi\right)$, $\tau_{FAST} = 0.125[s]$, $\tau_{SLOW} = 1.0[s]$
Z frequency weighting	No change to signal
A frequency weighting	zeros: $s = 4 \times 2\pi \cdot 0$ poles: $s = 2\pi\cdot$ $\quad 2 \times -20.599, 1 \times -107.653, 1 \times -737.862, 2 \times -12194.217$ $K = 7.5 \times 10^9$
C frequency weighting	zeros: $s = 2 \times 2\pi \cdot 0$ poles: $s = 2\pi\cdot$ $\quad 2 \times -20.599, 2 \times -12194.217$ $K = 5.9123 \times 10^9$
Sound Exposure, E	$E = \int_{t_1}^{t_2} p^2(t) dt$
Sound Exposure Level, SEL	$SEL = 10 \log_{10}\left(\frac{E}{E_0}\right), E_0 = p_0^2 t_0, p_0 = 20 \times 10^{-6}$ Pa, $t_0 = 1$ sec
Continuous-equivalent Level, LEQ	$LEQ = 10 \log_{10}\left(\frac{E}{p_0^2 T}\right)$

Table III: A list of commonly used Kernel machines, including the linear, polynomial, and Gaussian Radial Basis Function.

Kernel	$K(\mathbf{x}_i, \mathbf{z})$
Linear	$\mathbf{x}_i^T \mathbf{z}$
Gaussian Radial Basis Function (RBF)	$e^{-\gamma\left(\|\mathbf{x}_i - \mathbf{z}\|^2\right)}$

Table IV: The linear and RBF SVM accuracies for each experiment given a training set of 5000 feature vectors randomly sampled from the entire dataset. Shown are the average accuracies over 50 jackknifed samplings of the training set. Over the combined data set, the RBF kernel classifier performs 2.5% better than the linear classifier.

Experiment	ALL EXPS	APG1	APG2	Fort Drum	LRPE	MCBCL-PITT	SERDP-PITT	SERDP–CERL
Linear Accuracy (%)	91.3	94.8	92.2	82.8	91.0	90.3	92.4	94.7
RBF Accuracy (%)	93.8	96.3	95.5	86.5	92.9	93.3	92.9	95.3

Table V: The rank of each metric as obtained from the SVM-RFE algorithm, as well as the expected rank based on 40 rankings of different jackknife datasets.

Rank	Metric	Expected Rank
1	LZMAX slow	2.00
2	LZMAX fast	2.25
3	LCMAX fast	2.85
4	CLEQ	4.28
5	Zpeak	5.62
6	ZLEQ	5.93
7	LAMAX slow	6.60
8	ASEL	7.55
9	LCMAX slow	10.03
10	LAMAX fast	10.82
11	Apeak	12.03
12	Cpeak	12.12
13	ALEQ	12.35
14	ZSEL	12.45
15	CSEL	13.12

Table VI: The feature weights **w** when the top eight features are recursively selected from the full feature set, as well as the linear bias and the resultant blast/no-blast classifier accuracy, true positive rate (TPR), and false positive rate (FPR). Features that contribute more to a blast classification are shaded blue, features that contribute more to a non-blast classification are shaded red. Note: These weightings are relative to each jackknife sampling and are subject to vary for each random sampling.

LAMAX slow	ZLEQ	Zpeak	CLEQ	LCMAX fast	LZMAX fast	LZMAX slow	b	Accuracy	TPR	FPR
-0.0419	-0.2426	-0.1391	-0.1862	0.3604	0.2979	-0.0927	1.1138	0.9119	0.9084	0.0847
	-0.1547	-0.1567	-0.2546	0.3880	0.4065	-0.2659	0.7095	0.9077	0.9098	0.0944
		-0.1738	-0.3018	0.4341	0.4591	-0.4488	0.0177	0.9076	0.9111	0.0959
			-0.2349	0.2775	0.3631	-0.4267	-2.3707	0.8988	0.9022	0.1048
				0.0817	0.5582	-0.6304	-3.7287	0.8759	0.8617	0.1095
					0.7207	-0.7246	-3.9100	0.8557	0.8361	0.1242

List of Figures

List of Tables