

2009

# Analysis and Classification of Sounds Produced by Asian Elephants (*Elephas Maximus*)

Sharon Stuart Glaeser  
*Portland State University*

Follow this and additional works at: [https://pdxscholar.library.pdx.edu/open\\_access\\_etds](https://pdxscholar.library.pdx.edu/open_access_etds)



Part of the [Animal Sciences Commons](#), and the [Biology Commons](#)

Let us know how access to this document benefits you.

---

## Recommended Citation

Glaeser, Sharon Stuart, "Analysis and Classification of Sounds Produced by Asian Elephants (*Elephas Maximus*)" (2009). *Dissertations and Theses*. Paper 4066.  
<https://doi.org/10.15760/etd.5951>

This Thesis is brought to you for free and open access. It has been accepted for inclusion in Dissertations and Theses by an authorized administrator of PDXScholar. For more information, please contact [pdxscholar@pdx.edu](mailto:pdxscholar@pdx.edu).

THESIS APPROVAL

The abstract and thesis of Sharon Stuart Glaeser for the Master of Science in Biology were presented May 5, 2009, and accepted by the thesis committee and the department.

COMMITTEE APPROVALS:

[Redacted signature]

Randy Zelick, Chair

[Redacted signature]

Luis Ruedas

[Redacted signature]

Thomas Dolan

[Redacted signature]

David Shepherdson

[Redacted signature]

David Mellinger

DEPARTMENT APPROVAL:

[Redacted signature]

Michael Murphy, Chair  
Department of Biology

## ABSTRACT

An abstract of the thesis of Sharon Stuart Glaeser for the Master of Science in Biology presented May 5, 2009.

Title: Analysis and classification of sounds produced by Asian elephants (*Elephas maximus*)

Relatively little is known about the vocal repertoire of Asian elephants (*Elephas maximus*), and a categorization of basic call types and modifications of these call types by quantitative acoustic parameters is needed to examine acoustic variability within and among call types, to examine individuality, to determine communicative function of calls via playback, to compare species and populations, and to develop rigorous call recognition algorithms for monitoring populations.

This study defines an acoustic repertoire of Asian elephants based on acoustic parameters, compares repertoire usage among groups and individuals, and validates structural distinction among call types through comparison of manual and automated classification methods. Recordings were made of captive elephants at the Oregon Zoo in Portland, OR, USA, and of domesticated elephants in Thailand. Acoustic and behavioral data were collected in a variety of social contexts and environmental noise conditions. Calls were classified using perceptual aural cues

plus visual inspection of spectrograms, then acoustic features were measured, then automated classification was run. The final repertoire was defined by six basic call types (Bark, Roar, Rumble, Bark, Squeal, Squeal, and Trumpet), five call combinations and modifications with these basic calls forming their constituent parts (Roar-Rumble, Squeal-Squeak, Squeak train, Squeak-Bark, and Trumpet-Roar), and the Blow. Given the consistency of classifications results for calls from geographically and socially disparate subject groups, it seems possible that automated call detection algorithms could be developed for acoustic monitoring of Asian elephants.

ANALYSIS AND CLASSIFICATION OF SOUNDS PRODUCED BY  
ASIAN ELEPHANTS (*ELEPHAS MAXIMUS*)

by

SHARON STUART GLAESER

A thesis submitted in partial fulfillment of the  
requirements for the degree of

MASTER of SCIENCE  
in  
BIOLOGY

Portland State University  
2009

## **Dedication**

*This thesis is dedicated to my dear husband, Bobby, who offered his love and support throughout this adventure and encouraged me to share my heart with elephants.*

## **Acknowledgements**

First and foremost I would like to thank my thesis committee for their expertise and guidance: Dr. Randy Zelick, Dr. Luis Ruedas, Dr. Thomas Dolan, Dr. David Shepherdson, and Dr. David Mellinger. I am especially thankful to my advisor, Dr. Zelick, who met the challenges of data collection with enthusiasm and was so patient with me. I am eternally grateful to the Oregon Zoo elephant handlers, without whom this study would not have been possible. I especially appreciate the support from those most involved, April Yoder, Bob Lee, Joe Sebastiani, Jeb Barsh, and Dimas Dominguez. Dr. Mellinger and Dr. Holger Klinck, with the OSU/NOAA Cooperative Institute for Marine Resource Studies, provided invaluable supervision and support with acoustic analysis at a critical time. I appreciate the support and encouragement from the Oregon Zoo Conservation Program staff, Anne Warner, David Shepherdson, and Karen Lewis, and Deputy Director Mike Keele. I thank Dr. Mandy Cook for insight and guidance with early analysis, and Jean Lea Hofheimer and Bobbi Estabrook for scoring calls. I thank the managers at the Royal Elephant Kraal and Elephant Nature Park for the opportunity to record domesticated elephants in Thailand. Travel funding was granted by Portland State University Marie Brown and Academically-Controlled Auxiliary Activities Fund. I thank my loving family, Meghan Martin, and other dear friends who embraced this project with me. Finally, I am eternally grateful to Nancy Scott and the late Dr. LEL “Bets” Rasmussen for welcoming me into the elephant research community and paving the road, and to the many elephant handlers, veterinarians, and elephants who have taught me so much.

## TABLE OF CONTENTS

ACKNOWLEDGEMENTS .....	ii
LIST OF TABLES .....	v
LIST OF FIGURES .....	vi
CHAPTER I: INTRODUCTION .....	1
TAXONOMY AND CONSERVATION STATUS .....	2
ECOLOGY AND COMMUNICATION .....	4
VOCAL PRODUCTION AND INDIVIDUALITY IN ACOUSTIC SIGNALS .....	6
LOW-FREQUENCY SOUNDS .....	9
HEARING .....	11
PURPOSE AND SIGNIFICANCE OF THIS THESIS RESEARCH .....	12
CHAPTER II: PUTATIVE CALL TYPES AND MANUAL CALL CLASSIFICATION .....	16
METHODS .....	16
RESULTS .....	21
CHAPTER III: CALL DISTRIBUTION AND CALL RATE - HOW THE REPERTOIRE IS USED BY GROUPS AND INDIVIDUALS .....	27
METHODS .....	27
RESULTS .....	29
CHAPTER IV: ACOUSTIC PARAMETERS AND AUTOMATIC CLASSIFICATION OF CALLS .....	33
METHODS .....	33
RESULTS .....	53



CHAPTER V: DISCUSSION .....	78
REFERENCES .....	82
APPENDIX A: DISTRIBUTION PLOTS OF PARAMETERS.....	87
HISTOGRAMS AND BOXPLOTS OF ORIGINAL DATA .....	87
Q-Q PLOTS OF SCALED DATA .....	104
BOX PLOTS OF SCALED VARIABLES FOR IDENTIFYING OUTLIERS .....	111
APPENDIX B: CLASSIFICATION ON TREE RAW DATA OUTPUT .....	118
APPENDIX C: PRINCIPAL COMPONENT ANALYSIS DATA – RAW DATA OUTPUT .....	123
APPENDIX D: ANALYSIS OF SIMILARITY (ANOSIM) DATA OUTPUT .....	124
APPENDIX E: POWER SPECTRA OF RUMBLES AND NOISE.....	127

## LIST OF TABLES

Table 1: Acoustic ethogram of putative call types showing single calls, trains of calls, and call combinations.....	22
Table 2: Social contexts for call comparison.....	28
Table 3: Acoustic parameters measured by Osprey.....	38
Table 4: Calls in dataset for statistical analysis (N=927) .....	42
Table 5: Acoustic parameters that were used in statistical analysis.....	45
Table 6: Calls in dataset (N=908) after outliers removed.....	49
Table 7: Confusion matrix for tree.....	59
Table 8: Classification rate (%) for tree.....	59
Table 9: Summary of pair-wise ANOSIM showing significant difference between all call types.....	66
Table 10: Median values of acoustic parameters shown by the classification tree and PCA to differentiate basic call types.....	67
Table 11: Final ethogram for Bark, Roar, Rumble, Squeak, Squeal, Trumpet, and Blow.....	68

## LIST OF FIGURES

Figure 1: Call distribution of the Oregon Zoo herd versus domesticated elephants in Thailand .....	29
Figure 2: Call distribution of females in the Oregon Zoo herd. ....	30
Figure 3: Call distribution of one male (Tusko) in the Oregon Zoo herd. ....	31
Figure 4: Mean call rate as a function of social context.....	32
Figure 5: Osprey window of Squeak with annotation box and feature box. ....	34
Figure 6: Osprey window of Trumpet with annotation box and feature box. ....	35
Figure 7: Acoustic parameters measured by Osprey that can be visualized .....	37
Figure 8: Boxplot of scaled variables across all call types (blow, bark, roar, rumble, squeak, squeal, trumpet).....	43
Figure 9: Scaled variables for Roar showing outliers and gradients.....	48
Figure 10: Example histograms and boxplots of acoustic parameters for each call type show distribution and variability. ....	55
Figure 11: Final classification tree model .....	58
Figure 12: Principal Component Analysis plot for basic call types .....	61
Figure 13: Global ANOSIM test for difference in parameters among the call types...65	
Figure 14: Power spectra of Rumble and background noise at the Oregon Zoo.....	75
Figure 15: Power spectra of Rumble and background noise at Elephant Nature Park .....	76
Figure 16: Power spectra of rumbles and background noise at the Royal Elephant Kraal.....	77

## CHAPTER I: INTRODUCTION

Relatively little is known about the vocal repertoire of Asian elephants (*Elephas maximus*). Our knowledge of acoustic communication in elephants is based primarily on research on three wild African elephant (*Loxodonta spp.*) populations (Amboseli National Park in Kenya, Etosha National Park in Namibia, central African forests), three captive African elephant herds (Disney's Animal Kingdom in Lake Buena Vista, Florida, U.S.A., Vienna Zoo in Austria, and Daphne Sheldrick's orphanage in Nairobi National Park in Kenya), one wild Asian elephant population (Gal Oya National Park in Sri Lanka), and one captive Asian elephant herd (Oregon Zoo in Portland, OR, USA, formerly Washington Park Zoo). Additional elephant populations are currently under study, but publications are limited.

Over 30 calls have been published for African elephants (*Loxodonta spp.*) (Olson, 2004; Stoeger-Horwath et al., 2007), and nine have been published for Asian elephants (*Elephas maximus*) (Olson, 2004; McKay, 1973); but most are described only by context or function and not by acoustic features. Among the published African elephant calls, six infant calls have been described by both spectral and temporal features in addition to context and function, so there is advancement in our understanding of vocal ontogeny (Stoeger-Horwath et al., 2007). However, less than 16 calls produced by sub-adults and adults have been described by temporal or spectral features, and most of these appear to be variations of the low frequency rumble that differ in social context and function (Olson, 2004; Langbauer, 2000). Among the published Asian elephant calls, only the low frequency rumble has been

described by spectral and temporal features. African elephants are capable of producing calls ranging from 5 Hz to 9 kHz (Poole, in prep). Infrasonic calls (below 20 Hz) have been recorded from captive Asian elephants (Payne et al., 1986), but the frequency range of Asian elephant vocalizations is unknown.

#### **TAXONOMY AND CONSERVATION STATUS**

Elephants belong to the order Proboscidea and family Elephantidae. The two extant genera of the family Elephantidae include the African elephants (*Loxodonta africana*, Blumenbach 1769) and the Asian elephants (*Elephas maximus*, Linnaeus 1758).

Subspecies taxonomy of *Elephas maximus* varies among authors, but the International Union for Conservation of Nature and Natural Resources (IUCN) follows Shoshani and Eiesenberg (1982) in recognizing three subspecies: *E. M. indicus* on the Asian mainland (India and Indochina), *E. m. maximus* on Sri Lanka, and *E. m. sumatranus* on the Indonesian island of Sumatra (IUCN, 2008). These subspecies designations were based primarily on morphological differences, but mtDNA variation suggests that *E. m. sumatranus* is monophyletic (Fleischer et al., 2001). Borneo's elephants are not currently listed as a subspecies and their origin is controversial; however, comparison of mtDNA of Borneo elephants to that of Asian elephants across their range suggest that Borneo's elephants are genetically distinct, with divergence indicative of a Pleistocene colonization of Borneo (Fernando et al., 2003). Given genetic distinction, the IUCN could define these taxons as evolutionarily significant units (ESU) (IUCN, 2008).

The IUCN uses Sukumar's (2003) estimate of 41,410-52,345 wild elephants for the global population, but acknowledges the argument by Blake & Hedges (2004) that these estimates may be inaccurate due to the challenges of surveying forest-dwelling populations. A more recent estimate lists a wild population of 38,534-52,566 animals (Sukumar, 2006) and a global captive population of approximately 16,000 (Sukumar, 2006; Fischer, 2004). The Asian elephant is listed as Endangered by the IUCN (Choudhury et al., 2008. *Elephas maximus*. In: IUCN 2008) and is listed in Appendix I of the Convention on International Trade in Endangered Species of Wild Fauna and Flora (CITES) (CITES, 2008). The greatest threat to the Asian elephant is habitat loss, degradation, and fragmentation driven by an expanding human population, which leads to increasing conflicts between human and elephant. Poaching for ivory and other body parts also pose a major threat to the long-term survival of some populations. Large-scale hunting has reduced populations significantly over wide areas; but even selective removal of tusked males can result in skewed adult sex ratios and reduced genetic variation, especially in areas where populations are isolated (Choudhury et al., 2008. *Elephas maximus*. In: IUCN 2008).

Although there is genetic evidence that suggests there may be at least two species of African elephants, namely the African savanna elephant (*Loxodonta africana*) and the African forest elephant (*Loxodonta cyclotis*) (Roca et al., 2001), both the IUCN and CITES classify all African elephants as a single species encompassing both savanna and forest populations (IUCN, 2008; CITES, 2008). Many members of the international elephant management and medical community currently recognize

the reclassification of *Loxodonta africana* into two separate species, *Loxodonta africana* and *Loxodonta cyclotis*, with the following subspecies of *Loxodonta africana*, *L.a. africana* (South African bush elephant), *L.a. Knochenhaueri* (East African bush elephant), and *L.a. oxyotis* (West African bush elephant).

The African elephant is listed as Near Threatened by the IUCN (Blanc, 2008. *Loxodonta africana*. In: IUCN 2008) and is in Appendix I of CITES, with four populations down-listed to Appendix II (CITES, 2008). Poaching for ivory and meat remains a significant threat, but the most important perceived threat is habitat loss and fragmentation caused by an expanding human population and land conversion, which in turn leads to an increase in human-elephant conflict (Blanc, 2008. *Loxodonta africana*. In: IUCN 2008).

### ***ECOLOGY AND COMMUNICATION***

Elephants are intelligent, long-lived animals that live in a complex and fluid society in which several modes of communication play a role in maintaining group cohesion and social order, and in locating and assessing reproductive state of potential mates (Langbauer, 2000; Eisenberg et al., 1971). Asian and African elephants have a similar family structure. Females live in matriarchal herds consisting of other female relatives and their young. Elephant society is multi-tiered, with family units joining to form bond groups or larger clans that share a home range and coordinate movements (Douglas-Hamilton, 1972; Charif et al., 2005). Males are excluded from familial units at a young age and live as solitary elephants or in temporary association with other males in “bachelor herds” at the fringes of the female herds (Eisenberg et al., 1971;

Poole, 1987; Vidya & Sukumar, 2005). Recent studies suggest that bull society may be more structured than previously thought, and that young bulls may seek the company of mature bulls (Kate Evans, *pers. comms*). There is evidence that larger clans may not form in Asian elephant society, so ranging behavior may be influenced to a lesser degree by extended kin compared to African savannah elephants (Fowler & Mikota, 2006).

Data on Asian elephant mating behavior in the wild are scarce and incomplete, but studies of African savannah elephants suggest that mate choice by females and mate guarding by males appear to play crucial roles in the reproductive behavior of at least the African savannah species (Moss, 1983; Poole, 1989).

Infrasonic vocalizations and chemical signaling are considered primary modalities for long-distance communication (Payne et al., 1986; Poole, 1999); whereas visual, tactile, auditory, and chemical signals are all important at close range (Schulte & Rasmussen, 1999). Recent findings with seismic detection and discrimination suggest that vocalizations traveling in the seismic channel could be used for both short- and long-distance communication (O'Connell-Rodwell et al., 2006; O'Connell-Rodwell et al., 2007). Acoustic signals are temporally short-lived and both sender and receiver must be present for the signal to be communicated, and thus these signals provide information only on an immediate situation, for example, location (Langbauer, 2000). Chemical signals are temporally long-lived and the sender and receiver need not both be present for the signal to be communicated, and thus



these signals provide information on a constant state, for example, reproductive state (Langbauer, 2000).

### ***VOCAL PRODUCTION AND INDIVIDUALITY IN ACOUSTIC SIGNALS***

Mammals can modify certain basic sounds by changing the amplitude, temporal patterning, or stressing of overtones (Muckenhirn in McKay, 1973). Vocal individuality has been found in other mammalian species, for example bottlenose dolphins (*Tursiops truncatus*) (Janik et al., 2006), rhesus macaques (*Macaca mulatta*) (Rendall et al., 1998), fallow deer (*Dama dama*) bucks (Reby et al., 1998), swift fox (*Vulpes velox*) (Darden et al., 2003), but it is unknown how or if individual identity is communicated via the acoustic channel in Asian elephants. Structural variation within a call may serve a communicative function (Green, 1975 & Bayart et al., 1990 in Soltis et al., 2005), but call structure must first be characterized.

Characteristics of vocalizations that arise from inherent properties of vocal fold vibration in the larynx (the source) vary independently from properties that arise from vocal tract resonance (the filter), so either the source or filter may provide the auditory characteristics a receiver needs to determine identity (McComb et al., 2003). Air driven from the lungs sets the vocal folds in motion. Vibration of the vocal folds in the larynx has a *fundamental frequency* and *harmonics*. In terrestrial mammals, the fundamental frequency is determined by how many times the vocal folds vibrate in one second, measured in cycles per second or Hertz; and the fundamental frequency is negatively correlated to vocal cord mass (Fitch, 1997). The fundamental frequency determines the pitch that we hear. Harmonics of sound waves are integer multiples of

the fundamental frequency; that is, if the fundamental frequency is 25 Hz, then the harmonics have frequencies of 25 Hz, 50 Hz, 75 Hz, etc. The human auditory system does not perceive separate harmonics in sound, but harmonics provide the tone quality or richness of the sound.

The source sound from the larynx passes through the vocal tract, which selectively amplifies certain frequencies, and thus filters the spectral envelope to produce peaks called *formants* (McComb et al., 2002). As vocal tract changes shape, its resonance frequencies change, and different formants are produced. Articulators, the tongue and lips, produce more complex configurations of the vocal tract and hence more complicated formant patterns (Baken, 1996). The average spacing between consecutive formants, or formant dispersion, can provide information on the length of the vocal tract in mammals (McComb et al., 2003), with an increase in vocal tract length resulting in a decrease in formant spacing in many mammalian species (Sanvito et al., 2007; McComb et al., 2003; Fitch & Hauser, 2002).

The sound energy received is a factor of sound production intensity and acoustic propagation loss in the sound channel, which is influenced by topography, boundary conditions (surface and vegetation), and atmospheric structure (temperature, humidity, pressure) (Garstang, 2004).

The vocal tract of an elephant is the nasal passages of the skull, the trunk, and the pharyngeal pouch (Garstang, 2004). The pharyngeal pouch is at the posterior one-third of the nasopharynx (Fowler & Mikota, 2006). The elephant tongue is unable to protrude from the mouth due to the attachment from the tip of the tongue to the floor

of the mouth (Fowler & Mikota, 2006), so the tongue is limited in the articulation of sounds.

Adult Asian elephants have a size range of 2,000-5,500 kg (4,410-12,125 lbs) and 2-3.5 meters (6'7"-11'6") at the shoulder (Shoshani, 2000), and there is a visible difference between individuals in the size of the head and trunk, or the vocal tract. Center frequencies of formants depend on the size and posture of the caller, and different individual African savannah elephants have been found to overlap in these frequencies (McComb et al., 2002). However, there has been no comparative work done on the size of the vocal folds of elephants.

Soltis et al. 2005 (2005) found that structural variation in the rumble from adult female African elephants housed at Disney Animal Kingdom (Lake Buena Vista, Florida, U.S.A.) reflected individual identity and negative emotional arousal of caller. In the presence of dominant females, subordinate females produced rumbles with low tonality and unstable pitch compared to rumbles produced outside the presence of dominant females. Results of acoustic playback studies with African savannah elephants on female recognition of the contact call (McComb et al., 2003) and on signal assessment of the musth rumble (Poole, 1999) suggest individual- and size-related differences in acoustic production, which elicit questions regarding individualism, social recognition, and size assessment. Adult females were able to discriminate familiar and unfamiliar contact calls, and it appeared that the fundamental frequency contour extracted from the harmonics was the key characteristic in the signal that could be used for distinguishing individual calls at a distance (McComb et

al., 2000; McComb et al., 2003). Adult males responded to playback of musth rumbles in a way that suggests they are able to assess characteristics of the caller in this acoustic signal, but it is unclear whether they assess size in the signal or whether they recognize the caller. The musth rumble is a low-frequency vocalization produced only by males in musth, and is shown to advertise the musth state (Poole, 1987; Poole, 1989). The musth rumble has been studied extensively in the African elephant but not in the Asian elephant. In response to the musth rumble of high-ranking males, other musth males approached aggressively, while non-musth males walked away (Poole, 1999), which was consistent with prior studies of this population on the state of musth, dominance, and a willingness to contest access to females (Poole, 1989). Elephants are long-lived, intelligent animals, and they meet and interact over a period of decades, so it is reasonable to hypothesize that males also recognize individuals by their calls (Poole, 1999), and that they assess size based on prior interactions rather than assessing the size of an individual by their call. To further investigate the potential for size assessment in acoustic signals, the acoustic features that have the potential to code for size (fundamental frequency and formant dispersion) need to be measured for calls used over distances or barriers where visual access is limited.

### ***LOW-FREQUENCY SOUNDS***

Infrasound is an anthropocentric term for describing frequencies below the human hearing range, which is nominally 20 Hz to 20 kHz. The long wavelengths of low-frequency sound are more likely to hit only large objects and be reflected, and are thus resilient to atmospheric attenuation with distance (Langbauer et al., 1991; Larom

et al., 1997). Low-frequency sound below 100 Hz shows little attenuation in forest environments (Marter et al., 1977; von Muggenthaler et al., 2001). Infrasonic vocalizations of elephants have fundamental frequencies typically in the range 14 Hz to 35 Hz, with harmonics that extend into the audible range (Poole et al., 1988; Payne et al., 1986; Langbauer et al., 1989; Langbauer et al., 1991). Elephants can vocalize in the infrasonic range up to levels of 103 +/- 3 dB at 5m from the source (Garstang et al., 1995). The role and production mechanism of infrasound varies among species. Other species known to produce infrasonic sounds include the okapi, Bengal and Siberian tigers, giraffe, and Sumatran rhino (von Muggenthaler, 1992; von Muggenthaler, 2000; von Muggenthaler et al., 2001; von Muggenthaler et al., 2003).

The hyoid apparatus of elephants differs from the basic mammalian scheme and allows a greater flexibility of the larynx, which may aid in the production of infrasound (Langbauer, 2000). This difference in the hyoid apparatus presents a potential proximate cause for production of infrasound, and the need for long-distance communication presents an ultimate cause. McComb (2002) suggested that elephants produce fundamental frequencies in the infrasonic range simply because of their large size rather than an evolved mechanism for long-distance communication. McComb (2002) analyzed the source- and filter-related acoustic features of the African elephant female contact call in the Amboseli National Park population, measured degradation of the spectral structure of this call with distance, and performed playback to measure signal perception and assessment. Harmonics peaks around 115 kHz were most prominent and had the highest persistence with distances of 0.5 km to 2.5 km. Female

groups showed signs of detecting calls from distances of 2 km to 2.5 km, but did not respond as though they were familiar or unfamiliar until the distance narrowed to 1.0 or 1.5 km. These results suggest that although lower frequencies have the potential to travel further, the most important frequency components for long distance communication of social identity in African elephants may be in the harmonics around 115 Hz rather than in the infrasonic range of the fundamental frequency of the contact call (mean 16.8 Hz) (McComb et al., 2002).

### **HEARING**

There is some evidence that elephants may be better adapted for perceiving acoustic signals in the 100 Hz to 4 kHz range than in the range of the lowest frequencies produced. Heffner & Heffner (1982) measured the hearing sensitivity of a captive Asian elephant, and found the elephant to have an audibility curve similar to that of other mammals, but with a greater sensitivity to low frequencies and lower sensitivity to high frequencies than any other mammal tested prior to this study. The elephant's absolute threshold was 16 Hz (at 65 dB) to 12 kHz (at 72 dB), with a 17 Hz to 10.5 kHz hearing range by the 60 dB criterion. Although the elephant had a low frequency threshold of 16 Hz, it was considerably less sensitive to frequencies below 100 Hz than to those between 100 Hz and 4 kHz, and the maximum sensitivity was 1 kHz (at 8 dB). From 31.5 Hz to 2 kHz the thresholds were close to background noise level, so it is possible that the animal's actual sensitivity in this region was masked by background noise. Frequency discrimination tests indicated the elephant's frequency discrimination was best below 1 kHz.

Localization of sound relies on the time of arrival to the ear and the intensity of sound reaching the ear. Because the interaural distance in the elephant is large, sound reaches one ear long before the other, and the intensity of the sound on the ear closest to the sound is greater than on the opposite ear due to the buffering effects of the head. The larger the interaural distance, the greater the ability to localize low frequency sounds because the greater interaural distance means a larger phase and amplitude and difference of the incoming sound wave between the two ears. Heffner & Heffner (1982) conducted localization experiments for a frequency range of 125 Hz to 8 kHz, source angles of 0° to 60°, and various stimuli. The elephant performed best at localizing sound below 300 Hz and was virtually unable to localize 4 kHz to 8 kHz. Pinnae extension appeared to play a role in localizing sound in this study (Heffner et al., 1982).

### ***PURPOSE AND SIGNIFICANCE OF THIS THESIS RESEARCH***

The purpose of this research is to contribute to the basic science of elephant communication and to conservation efforts. The specific goals are to (1) define an acoustic repertoire of Asian elephants based on acoustic parameters, (2) investigate how the repertoire is used by groups and individuals, (3) compare manual and automated classification to validate structural distinction among call types, (4) provide a basis for comparing acoustic communication among elephant species and populations, and (5) explore the potential for using call parameters to develop an automatic detector of Asian elephant calls for acoustic monitoring applications.

Communication is the intended transfer of information (a signal) from a sender to either an intended or unintended receiver. Communication is intended to confer some advantage to the sender, receiver or both via natural or sexual selection. This study analyzes sounds produced by elephants, but there was no examination of an intended transfer of information by measuring the response of conspecifics, so it is not implied that these sounds are signals with communicative value. Data were collected for future investigation of the communicative function of these sounds, but this topic was not included in this analysis.

This study provides a basis for future research. A categorization of basic call types and modifications of these call types by quantitative acoustic parameters is needed to examine acoustic variability within and among call types, to examine individuality, to determine communicative function of calls via playback, to compare species and populations, and to develop rigorous call recognition algorithms for monitoring wild and managed populations.

The task of vocalization classification and speaker identification is common in bioacoustic analysis. Automated methods are particularly useful with large datasets or when the repertoire is being compared between individuals or social groups (Deecke & Janik, 2006). Detection algorithms are used extensively in the passive acoustic monitoring of marine mammals, and there are many published classification methods and tools that could be used with other vocalization datasets. Examples of methods referenced for this current study include classification of African elephant and marine mammals vocalizations. Campbell et al. (2002) used artificial neural network to



identify individual female Steller sea lions (*Eumetopias jubatus*). Clemins et al. (2005) developed a Hidden Markov Model based on human speech recognition techniques to automatically classify African elephant vocalizations for call type and individual.

Classification of sounds by acoustic parameters is the first step in developing call recognition algorithms for call detection and acoustic monitoring or census. Payne et al. (2003) provides evidence that elephant calling patterns can be reliable indicators of group size and composition, both of which are important for acoustic monitoring. Calls were divided into three structures, single-caller low-frequency, single-caller high frequency, and multiple-callers low-frequency. The rate of calling increased with increasing numbers of elephants, and the distribution of these call categories changed with group composition. Vocalizations may provide another tool for mitigating human-elephant conflict, which often results in injury or death to both humans and elephants (Kemf & Santiapillai, 2000). In Sri Lanka, researchers are investigating the use of speech recognition techniques as a means of remotely identifying individual elephants as they approach crops (Doluweera et al., 2003).

A library of vocalizations from known individuals in known reproductive states will facilitate a more rigorous categorization of sounds by acoustic parameters, and will allow an examination of sources of variability. An understanding of this variability is needed to reliably determine the meaning of various calls via playback (Langbauer, 2000). A database of acoustic communication of African savannah elephants (*Loxodonta africana*) is currently being developed by the Savannah

Elephant Vocalization Project (SEVP) in Amboseli National Park (SEVP, 2008) and by Disney's Animal Kingdom (John Lehnhardt, Joseph Soltis, pers comms). SEVP has collected more than 70 different elephant call types of African savannah elephants and linked them to observations of elephant behavior (SEVP, 2008), but the descriptions of only a subset of these calls have been published. There is no similar database known to be developed for wild or captive Asian elephants (*Elephas maximus*), so one goal of this current study is to contribute to an animal communications database for future research.

This study aims to provide a basis for comparisons of acoustic communication in wild and captive Asian elephants, and could potentially serve as a basis for comparisons between Asian and African elephants. These comparisons may provide insights into the ecological role such communication plays and the requirements of counterpart populations. As elephants lose habitat and find themselves under varying degrees of management, the need to understand their requirements in captivity will become even more important to the survival of the species (Riddle et al., 2003).

## CHAPTER II: PUTATIVE CALL TYPES AND MANUAL CALL CLASSIFICATION

### *METHODS*

#### **Study site and study subjects**

Data were collected at the Oregon Zoo in Portland, Oregon, U.S.A, and in Thailand. Data collected at the Oregon Zoo were from a captive community of three male and four female Asian elephants. During the course of this study, the females were managed as a single herd with only temporary separations. The males were housed alone with no visual or physical access to other males; however, they did have acoustic and olfactory access given their close proximity and the movement of elephants between enclosures. Males and females were housed together during social introductions and breeding. The Oregon Zoo elephant exhibit is designed with two outside yards totaling 3140 m<sup>2</sup>, and seven inside rooms. The Oregon Zoo is in an urban setting located near a major interstate. Sources of anthropogenic noise include highway traffic, air traffic, hydraulics, water, electric fences, zoo construction, a zoo train, and visitors.

Data collected in Thailand were from two herds of domesticated elephants, one herd of approximately 80 elephants at the Royal Elephant Kraal in urban Ayutthaya, and one herd of approximately 30 elephants at the Elephant Nature Park in the rural Mae Taeng Valley north of Chiang Mai. The Royal Elephant Kraal is a working elephant village with adult and semi-adult bulls and cows, geriatric cows, and a nursery of 10-12 calves. The Elephant Nature Park manages injured and abused

animals in a semi-captive setting where elephants are allowed to roam during the day within the grounds and chained at night in social groups. At the time of recording, there were three adult bulls, two sub-adult males, two calves, and females of varying ages. Sources of anthropogenic noise in both Thailand sites were sporadic road noise, machinery, water, and faint talking of visitors.

### **Acoustic data collection**

Data were collected at the Oregon Zoo from February 2005 to March 2009, and consisted of over 80 hours of observations and recordings, with 56 hours of usable data. Data were collected continuously for one hour (occasionally 30 min to 2 hours) during social introductions, breeding events, temporary separations, arrival of a bull, the death of a matriarch, novel events, and routine husbandry. The following data were collected during each session: social context, behavior, vocalizations, and visual signs of musth. Reproductive state measured by hormone levels was provided by the Oregon Zoo. The caller, if identified during observation by sound localization or visual cues, was dictated into the recorder or noted on a checksheet. One challenge was detecting and localizing vocalizations at low frequency; however the harmonics often extended into the audible range, making it possible to detect and localize at close proximity to the caller .

Data were collected in Thailand from November to December 2009, and consisted of approximately 6 hours of data. Data were collected continuously for one hour sessions (or opportunistically for short sessions) during the morning release to pasture, morning routines, night feedings, greetings between human and elephant, and

elephant painting. The following data were collected during each session: social context, vocalizations, and focal animal if there was a focal.

Acoustic data were collected with four systems during the course of the study. The frequency responses given for these systems are those provided by the manufacturer, unless otherwise specified. The systems used were: (1) M-Audio MicroTrack II (20 Hz – 20 kHz, +/- 0.5 dB, recorder tested down to 10 Hz +/- 3 dB), (2) Bruel & Kjaer 4145 condenser microphone (2.6 Hz -18 kHz +/- 2 dB), an ACO type 012 preamp (flat to 0.5Hz), an ACO PS2000 power supply, and a Racal V-Store 24 Instrumentation Recorder (analog, DC-45.5 kHz), (3) Edirol R-09 mp3 recorder (20 Hz – 40 kHz +/- 2 dB), (4) acoustic extraction from digital video (Panasonic PV-DV700, 20 Hz – 20 kHz) that is being used as part of an ongoing behavioral study. Only the M-Audio and digital video were used to collect data in Thailand. Preliminary data were collected with an IoTech Wavebook 512 12-bit 1MHz Data Acquisition System, USBGear USB Sound card modified to record to DC, and Vetter 820 analog recorder.

During recording sessions, the microphone was fixed in position with a tripod to minimize extraneous noise from holding it directly, as per McComb (1996). The distance to source was variable as subjects moved throughout the recording session, so distance to source was only approximated for each recording session and there was no attempt to measure absolute intensity of calls.

A Cambridge Electronic Design (CED) Micro1401 mkII data acquisition unit running Spike2 software (v5.12) was used to upload analog acoustic data. Adobe

Audition (v1.5) was used to convert file formats and split recording channels. Files were first down-sampled to 25kHz to reduce the file size. The sampling rate was selected based on the published frequency range of African elephant calls, 5 Hz to 9 kHz (Poole, in prep) which gives a minimum sampling rate at the Nyquist frequency of 18 kHz. However, energy was found to extend above 12 kHz, so the original sampling rate of either 32 kHz or 44 kHz was used for final measurements of acoustic parameters.

### **Scoring calls and classifying calls into putative call types**

The ethogram used for preliminary data collection included vocalizations described by McKay (1973) and those in the Elephant Husbandry Resource Guide (Olson, 2004). The ethogram of elephant behaviors in the Elephant Husbandry Resource Guide is a compilation of ethograms from approximately 30 publications and manuscripts, so it is a relatively comprehensive source. Ad libitum sampling was employed to incorporate vocalizations not yet described. Six basic call types were defined (trumpets, squeaks, squeals, roars, rumbles, and barks), and these call types and modifications of these call types were fit into published nomenclature where possible. Combination calls were also defined (trumpet-rumble, trumpet-roar, squeak-rumble, squeak-squeal, roar-rumble). Call descriptions were compared to those recorded in an ongoing study in Uda Walawe National Park in southern Sri Lanka, and a joint ethogram was developed for calls that were common to the captive herd at the Oregon Zoo and the free-ranging elephants in Sri Lanka (Shermin de Silva, pers comms.)

Exemplar calls were presented to elephant handlers and university students to gather descriptions of how the sounds were perceived compared to familiar non-elephant sounds, with the aim of providing a description that would allow elephant handlers and researchers to discuss elephant calls in terms of aural cues; for example, a Squeal sounds like rubber shoes on a wood floor.

A signal or sound can be viewed as a distribution of energy in time and frequency. These distributions can be represented several ways. A *spectrogram* displays a three-dimensional plot that shows how the sound varies over time, with frequency on the y-axis, time on the x-axis, and the relative power (or the logarithm of relative power) at a given point in frequency and time represented as a darkness value. A *power spectrum* displays average energy of the signal over a period of time, with relative power on the y-axis and frequency on the x-axis (frequency domain). A *waveform* displays amplitude on y-axis and time on x-axis (time domain).

Praat software (v4.5.16, Institute of Phonetic Sciences of the University of Amsterdam, The Netherlands) was used to annotate (score) sounds by adding boundary markers and typing call notations (call type, caller ID, quality) in an annotation field within the boundary. Calls were categorized into the putative call types using perceptual aural cues and visual inspection of spectrograms for differentiation of the following acoustic parameters: fundamental frequency contour (start and end frequency, maximum and minimum frequency, inflection), tonality, and signal duration. Categorizing by aural cues and visual inspection is considered a manual method of classification. With the exception of blow sounds, every detectable

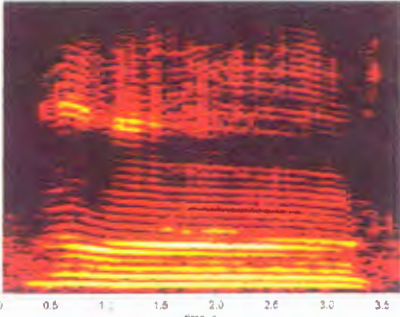
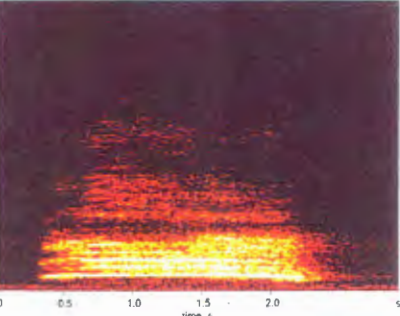
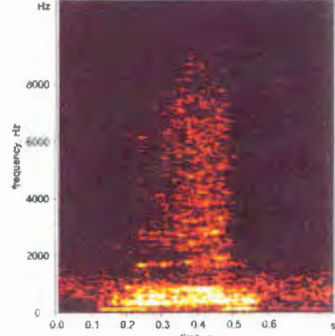
occurrence of every call type was scored. Environmental sounds were also scored for future analyses with call detection. Half of the recordings were first scored by two trained but inexperienced reviewers, then verified and corrected. Of 1243 sounds scored by inexperienced reviewers using aural cues and spectrogram inspection, 90 required correction, for an interobserver reliability of 93%. The caller (if known) was annotated for each call. A quality score of 1 to 4 was assigned to each call using methods by Cambell et al. (2002), which were based on presence/absence of sporadic sound overlap (elephant, human, environmental) and a subjective measurement of degradation by ambient noise (e.g., highway, wind, water, electricity, faint sounds of visitors, insects).

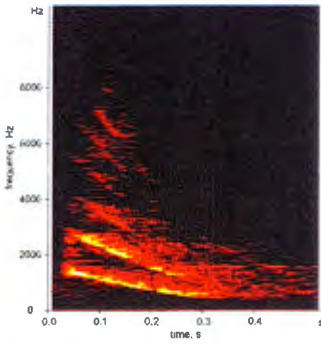
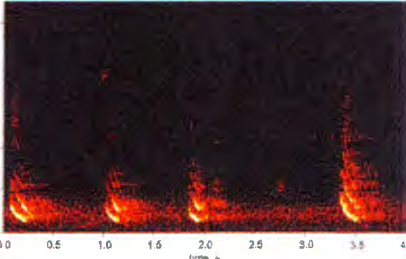
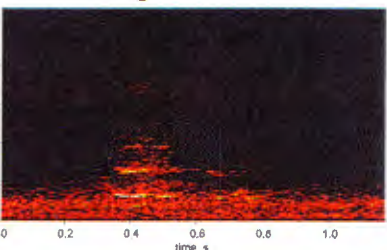
## ***RESULTS***

The ethogram of putative call types with descriptions and exemplar spectrograms is provided in Table 1. The aural cue and spectrogram descriptions were used for categorizing sounds into call types. The visual cue was used to help identify the caller. The published descriptions include only those for Asian elephants.

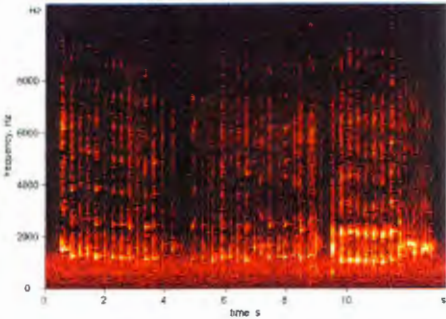
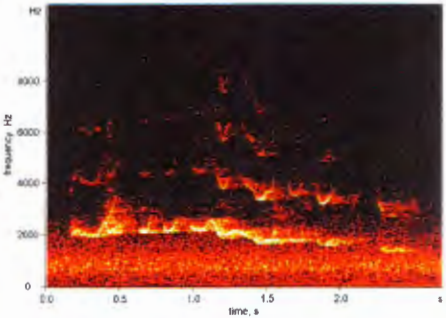
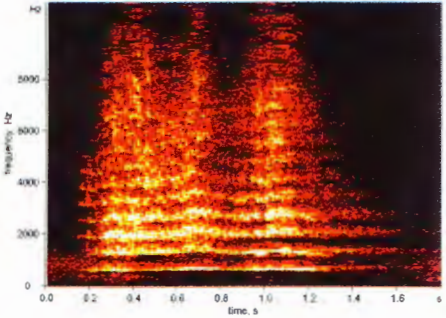
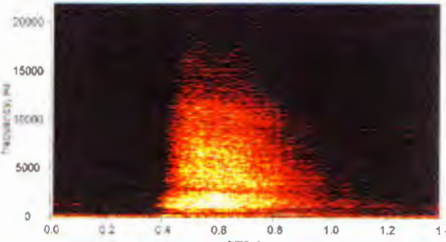


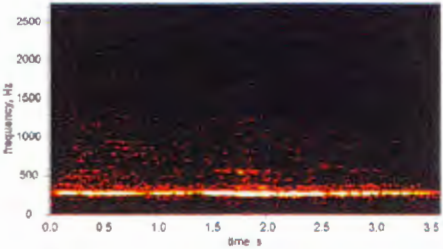
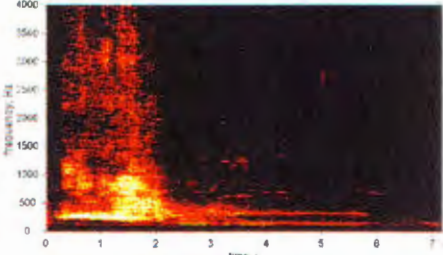
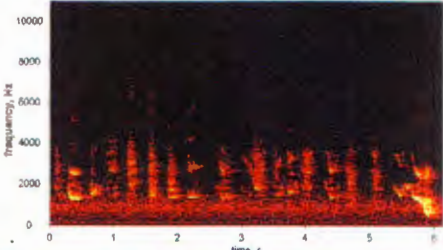
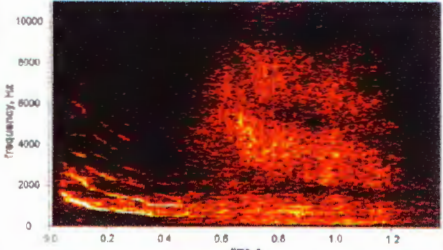
**Table 1: Acoustic ethogram of putative call types showing single calls, trains of calls, and call combinations.**

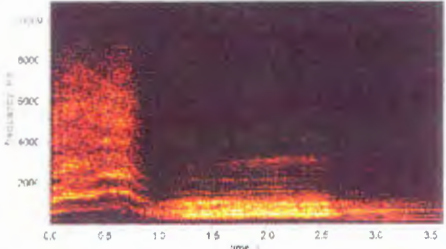
Call type and spectrogram	Definition
<p data-bbox="277 401 400 430"><b>RUMBLE</b></p> 	<p data-bbox="593 354 1192 439"><i>Aural cue:</i> Sounds like a diesel engine, motorboat, distant helicopter. Pulsating with a flat pitch. Perceived as low frequency.</p> <p data-bbox="593 458 1165 573"><i>Spectrogram inspection:</i> Flat fundamental frequency with narrow bandwidth. Pulsating in spectrogram associated w/ amplitude modulation in waveform. Harmonics present.</p> <p data-bbox="593 592 1139 639"><i>Visual cue of caller:</i> Mouth visibly open, forehead flutter.</p> <p data-bbox="593 658 1192 801"><i>Published description:</i> A resonant growl (Olson, 2004). Growl modified by resonance in trunk, used in context of mild arousal (McKay, 1973). Call with fundamental frequency in range 14 Hz to 30 Hz, lasting 10-15 sec.</p>
<p data-bbox="297 906 380 934"><b>ROAR</b></p> 	<p data-bbox="593 830 1179 915"><i>Aural cue:</i> Sounds like a lion roar with changing pitch. Pitch sounds like it changes because of amplitude modulation. Perceived as low frequency.</p> <p data-bbox="593 934 1152 1106"><i>Spectrogram inspection:</i> Flat or slightly modulated fundamental frequency. Amplitude modulated, with greatest amplitude somewhere during the call rather than at beginning or end. Tonal, noisy, or mixed. Sometimes ends in rumble. Broader bandwidth than rumbles, but inconsistent. Harmonics present.</p> <p data-bbox="593 1125 1119 1172"><i>Visual cue of caller:</i> Mouth wide open or widens during call, lower abdomen contracts.</p> <p data-bbox="593 1191 1192 1334"><i>Published description:</i> Growl modified by amplitude increase, used in context of long-distance contact. Used primarily by juveniles who have been separated from their group (McKay, 1973). Pulsating sound. Loud growl (Olson, 2004).</p>
<p data-bbox="297 1353 380 1382"><b>BARK</b></p> 	<p data-bbox="593 1382 1159 1467"><i>Aural cue:</i> Sounds like a short loud grunt, clearing throat, cape buffalo, chuff of cow, cheetah. Ends abruptly with no roll off. Noisy compared to rumble.</p> <p data-bbox="593 1487 1165 1572"><i>Spectrogram inspection:</i> Flat fundamental frequency. Short duration. Call is usually noisy. Harmonics not always visible (or difficult to see with noise).</p> <p data-bbox="593 1591 955 1620"><i>Visual cue of caller:</i> Mouth open.</p> <p data-bbox="593 1639 1165 1705"><i>Published description:</i> None for Asian elephants, but may be similar to the bark African elephant calves (Stoeger-Horwath et al., 2007).</p>

<i>Call type and spectrogram</i>	<i>Definition</i>
<p style="text-align: center;"><b>SQUEAK</b></p> 	<p><i>Aural cue:</i> Sounds like a wet finger across tight rubber surface, audible pedestrian crossing signal, baby alligator, gibbon call. High-pitched sound that falls in pitch, powerful initial sound.</p> <p><i>Spectrogram inspection:</i> Descending fundamental frequency that is either straight or curved in shape. Small short blip may precede pitch fall. Harmonics present. Modified by repetition.</p> <p><i>Visual cue of caller:</i> Cheeks depressed, mouth open slightly or closed, tail sometimes erect.</p> <p><i>Published description:</i> Not heard in wild (McKay, 1973).</p>
<p style="text-align: center;"><b>SQUEAK TRAIN</b></p> 	<p><i>Aural cue:</i> Train of repeated Squeaks with short breaks between Squeak sounds of about 1 sec. Rate of squeaking may change.</p> <p><i>Spectrogram inspection:</i> Squeaks are in a train if time between edges (end of one/beginning of next) of Squeaks is less than 1 sec. (It was easier to perceive a break of 1 sec than 0.5 sec, so 1 sec was used to define a train versus a bout of individual calls.) Amplitude sometimes shows decrease of subsequent Squeaks towards the end of the train. Harmonics present.</p> <p><i>Visual cue of caller:</i> Cheeks depressed, mouth open slightly or closed, tail sometimes erect.</p> <p><i>Published description:</i> Chirp= multiple short squeaks. Possibly in context of conflict (McKay, 1973).</p>
<p style="text-align: center;"><b>SQUEAL</b></p> 	<p>Sounds like rubber shoes on a gym floor, release of air out of a balloon, windshield wiper on a dry window, a dog whine. Short utterance.</p> <p><i>Acoustic description:</i> Flat or modulated pitch. Short duration (less 1-1.5sec from aural cues), usually flat in pitch or difficult to tell (because it looks like just a point). Harmonics present. Modified by repetition.</p> <p><i>Visual cue of caller:</i> Cheeks depressed, mouth open slightly or closed.</p> <p><i>Published description:</i> none</p>



<i>Call type and spectrogram</i>	<i>Definition</i>
<p style="text-align: center;"><b>SQUEAL TRAIN</b></p> 	<p><i>Aural cue:</i> Train of repeated Squeal with short breaks between sounds. Sounds like squeals in staccato, intermittent release of air from a balloon. Each squeal sounds flat in pitch but they jump up and down.</p> <p><i>Spectrogram inspection:</i> Squeals are in a train if time between edges (end of one/beginning of next) of Squeals is less than 1 sec. Flat fundamental frequency of each squeal, with frequency shifts or jumps between Squeals in the train. Harmonics present.</p> <p><i>Visual cue of caller:</i> Cheeks depressed, mouth open slightly or closed.</p> <p><i>Published description:</i> none</p>
<p style="text-align: center;"><b>SQUEAL LONG</b></p> 	<p><i>Aural cue:</i> Sounds like continual release of air from a balloon, fingers on a wet balloon, sliding pitch, sounds long.</p> <p><i>Spectrogram inspection:</i> Modulated fundamental frequency with many inflection points. Longer than Squeal but duration variable, clearly longer than 1.5-2sec. Sound is continuous with no apparent breaks and pitch slides rather than jumps. Frequency values similar to Squeal Trains but with continuous energy rather than breaks. Harmonics present.</p> <p><i>Visual cue of caller:</i> Cheeks depressed, mouth open slightly or closed.</p> <p><i>Published description:</i> none</p>
<p style="text-align: center;"><b>TRUMPET</b></p> 	<p><i>Aural cue:</i> Sounds like a trumpet blast. Perceived as a higher frequency call.</p> <p><i>Spectrogram inspection:</i> Flat fundamental frequency. Narrowband. Sometimes has broadband energy at higher frequencies (top hat of noise). Harmonics present.</p> <p><i>Visual cue of caller:</i> maybe a lifting at base of trunk, trunk sometimes extended, tail sometimes erect sometimes (any behavior indicating extreme arousal).</p> <p><i>Published description:</i> Squeak modified by increased amplitude and duration of call. Used in context of extreme arousal. Pulsating sound (McKay, 1973).</p>
<p style="text-align: center;"><b>BLOW</b></p> 	<p><i>Aural cue:</i> Forced exhalation through end of trunk. Not a vocalization.</p> <p><i>Spectrogram inspection:</i> Noisy broadband. Looks like a foot scrape, but lower frequency than foot scrape. Harmonics not clearly visible.</p> <p><i>Visual cue of caller:</i> air blast from trunk</p> <p><i>Published description:</i> May be same as snort in McKay (1973).</p>

Call type and spectrogram	Definition
<p style="text-align: center;"><b>BLOW-TONAL</b></p> 	<p><i>Aural cue:</i> Sounds like blowing into a glass bottle or rubbing water on the rim of a crystal glass. Very clear.</p> <p><i>Spectrogram inspection:</i> Tonal. Flat fundamental frequency. Harmonics not clearly visible.</p> <p><i>Visual cue of caller:</i> Forehead protrudes enough to see at a distance. Produced only by one animal in study (Oregon Zoo male, Tusko).</p> <p><i>Published description:</i> none</p>
<p><b>Combination calls</b></p> <p>Calls in which constituent parts are basic calls types, but with a continuous frequency contour and no apparent break in sound between the constituent parts. The criterion of continuity differentiate combination calls produced by a single individual from overlapping calls produced by multiple individuals.</p>	
<p style="text-align: center;"><b>ROAR/BARK + RUMBLE</b></p> 	<p>A Roar or Bark ending in a Rumble.</p> <p>Same acoustic structure as singles calls, but with continuous sound and frequency contour.</p> <p>May just be a variance of Roar. Many Roars have at least a slight fall off at the end that sounds like a Rumble. In (Stoeger-Horwath et al., 2007), Roars are broken down into mixed Roars of various combinations.</p>
<p style="text-align: center;"><b>SQUEAL + SQUEAK</b></p> 	<p>Any combination of Squeals and Squeaks occurring in a train with less than 1 sec between the edges of the calls.</p> <p>Same acoustic structure as singles calls.</p>
<p style="text-align: center;"><b>SQUEAK + BARK</b></p> 	<p>A Squeak ending with any variation of Bark, Rumble, or Roar. Usually ends with a Bark.</p> <p>Same acoustic structure as singles calls, but with continuous sound and frequency contour.</p>

<i>Call type and spectrogram</i>	<i>Definition</i>
<p><b>TRUMPET + ROAR/RUMBLE</b></p> 	<p>A Trumpet ending in a Roar or Rumble. Trumpet ending in Rumble is rare.</p> <p>Same acoustic structure as singles calls, but with continuous sound and frequency contour.</p>

## **CHAPTER III: CALL DISTRIBUTION AND CALL RATE - HOW THE REPERTOIRE IS USED BY GROUPS AND INDIVIDUALS**

### ***METHODS***

#### **Comparing call distribution of groups and individuals**

To compare incidence of call types among groups and individuals, the number of calls of each putative call type (Table 1) was totaled for the study sites and for individuals in the Oregon Zoo herd. For this comparison, the following calls were removed from the dataset: duplicates of calls recorded with multiple systems, Blows, and recordings of one female elephant who vocalized on command. All combination calls with basic call types as the constituent parts were combined into one group called “combo.”

The call rate was not weighted to account for differences in recording time among sessions. For comparison of the Oregon Zoo herd to domesticated elephants in Thailand, the recording duration, sample size of calls, and number of subjects were very different, but the contexts in both sites offered situations of mild to extreme arousal. For comparing individuals within the Oregon Zoo herd, the call rate was not weighted because the total recording time for the cows was similar. Also, the caller has not yet been identified for every call, so these call compositions provide only a snapshot of how each cow uses the repertoire. Only one male, Tusko, was included in this comparison because of limited data for the other males.

## Comparing call rate in various social contexts

The rate of calling, regardless of the caller, was compared among the various contexts in which data were collected at the Oregon Zoo. For this comparison, the following calls were removed from the dataset: duplicates of calls recorded with multiple systems, Blows, recordings of one female elephant who vocalized on command, and males recorded when by themselves. The total number of calls was 2147. The number of calls of each call type per recording session were calculated, then adjusted to a 60 minute recording duration. The social contexts in which recordings were made are shown in Table 2. Most of the recordings were scheduled during social introductions, with some opportunistic recordings.

**Table 2: Social contexts for call comparison**

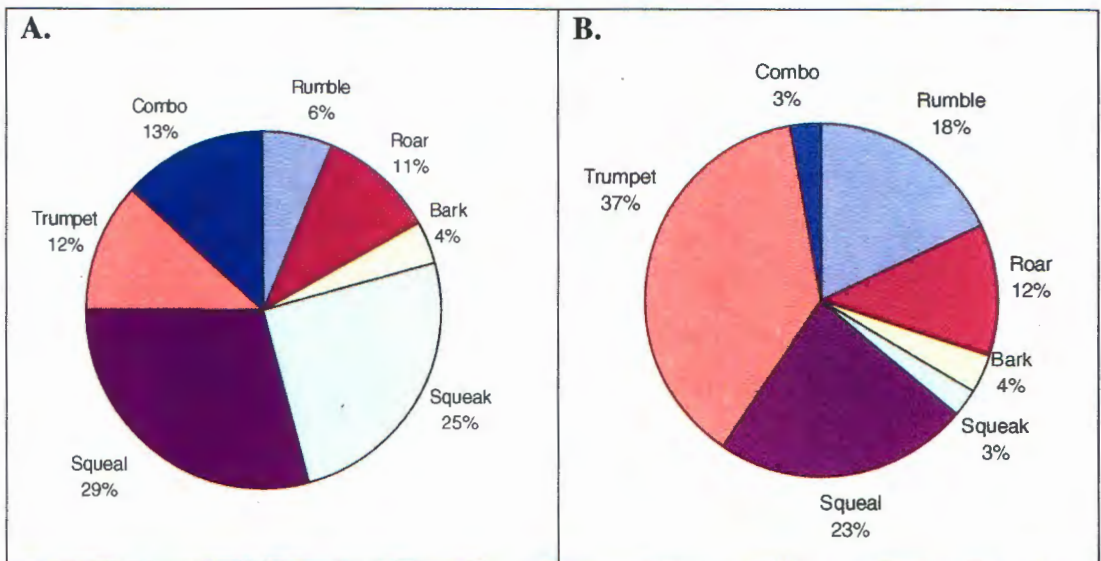
Context	Description	# of sessions
Routine	No specific event, normal training, husbandry, shifting	6
Enrichment	Novel enrichment in yard (not just browse, maybe new tires or something rare)	3
Introduction	1 male in with female group (breeding, intro), includes howdy	39
Reunion after introduction	Females join after separation of group for an Introduction event	1
Separation	Focal animal is separated and by themselves (not with male as in intro)	2
Reunion after separation	Females join after a Separation event	1
Transfer	Arrival of bull Tusko	1
Death	Euthanasia of matriarch Pet, including time during euthanasia, visitation of herd members to the body, other herd members after leaving the body	1



## RESULTS

### Call distribution between sites

With the exception of the Squeak call, the domesticated elephants recorded in Thailand produced the same repertoire as the Oregon Zoo elephants, but with a different distribution of calls, as shown in Figure 1. The Squeak call was made by Thai elephants, but only when asked to speak or when engaging tourists. Some factors that potentially account for the difference in call distribution are a difference in social context, difference in number of animals in the herd, environmental noise, and individuality.



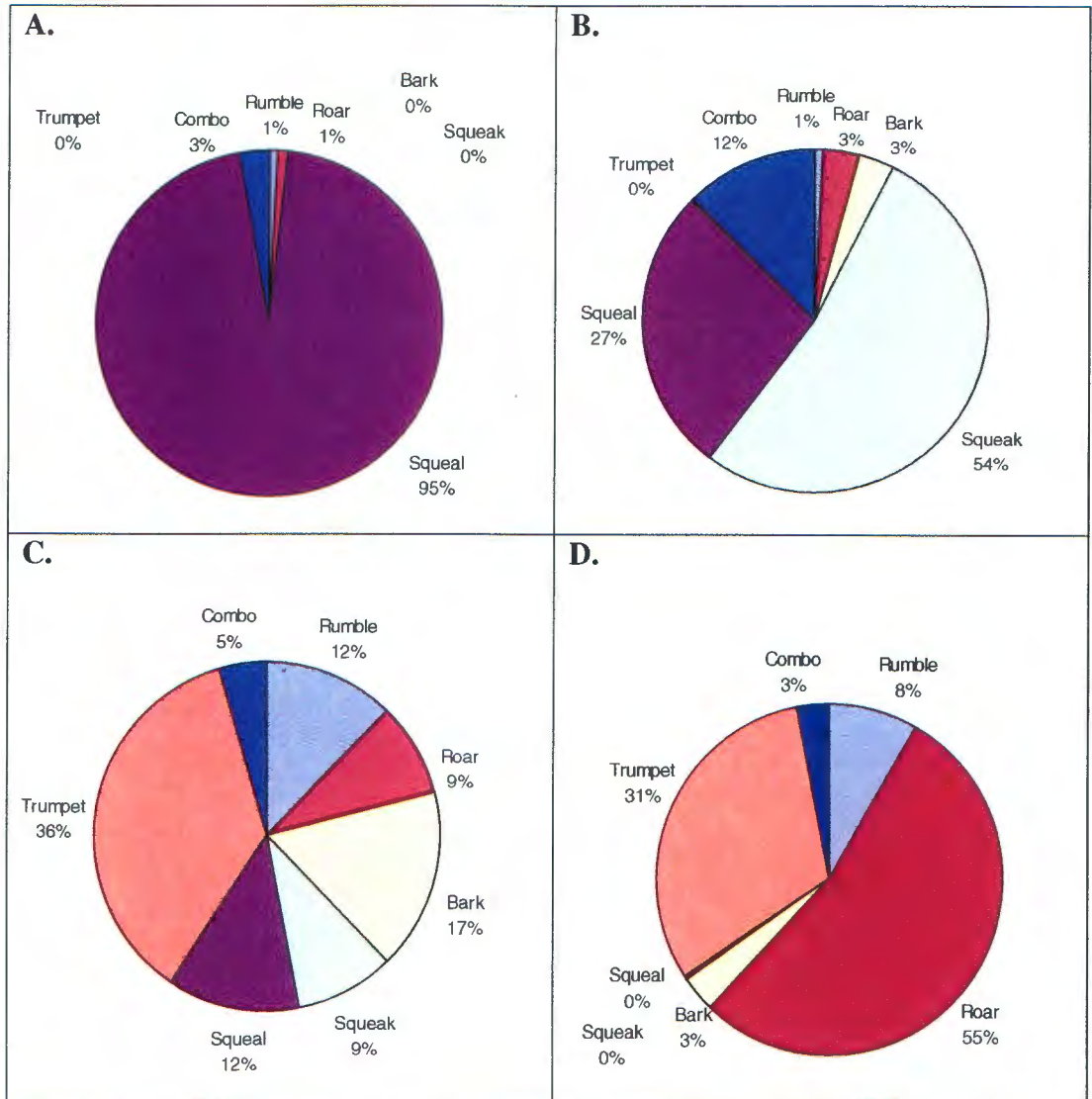
**Figure 1: Call distribution of the Oregon Zoo herd versus domesticated elephants in Thailand**  
Panel A shows the call distribution of the Oregon Zoo herd (N=2066 calls). Panel B shows the call distribution of domesticated elephants in Thailand (N=279 calls).

### Call distribution among individuals

Individuals within the Oregon Zoo herd use the repertoire quite differently, as shown in Figure 2 and Figure 3. For most of the recording sessions, the cows were together as a group, so the difference in calls produced could be a function of

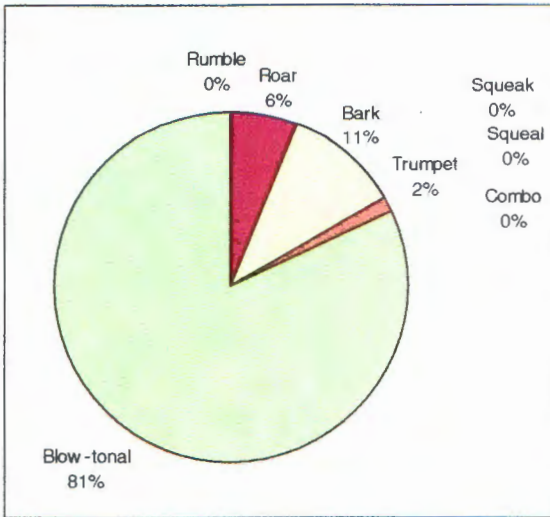


individuality, differences in roles within the herd, or differences in level of arousal within the same social context. Investigating these differences is planned with future analyses. Only one of the three bulls was recorded consistently during the course of this study, and this bull produced primarily one sound, a tonal Blow.



**Figure 2: Call distribution of females in the Oregon Zoo herd.**

Panel A shows Pet. Panel B shows Sung-Surin (Pet's daughter). Panel C shows Rose Tu. Panel D shows Chendra. Panel E shows Tusko (male).



**Figure 3: Call distribution of one male (Tusko) in the Oregon Zoo herd.**

### **Call rate in various social contexts**

Figure 4 shows a comparison of call rates as a function of social context. The highest rate was in temporary separations and reunion after those separations. Not surprising is the low rate of calling during routine husbandry. What is surprising is the relatively low rate of calling during the death of the matriarch, the transfer in of a bull, and social introductions. Given that elephants use many modes of communication, it is possible that the rate of acoustic production for this herd is not a reliable indicator of arousal, or that what we perceive as a situation that warrants increased arousal does not actually do so.

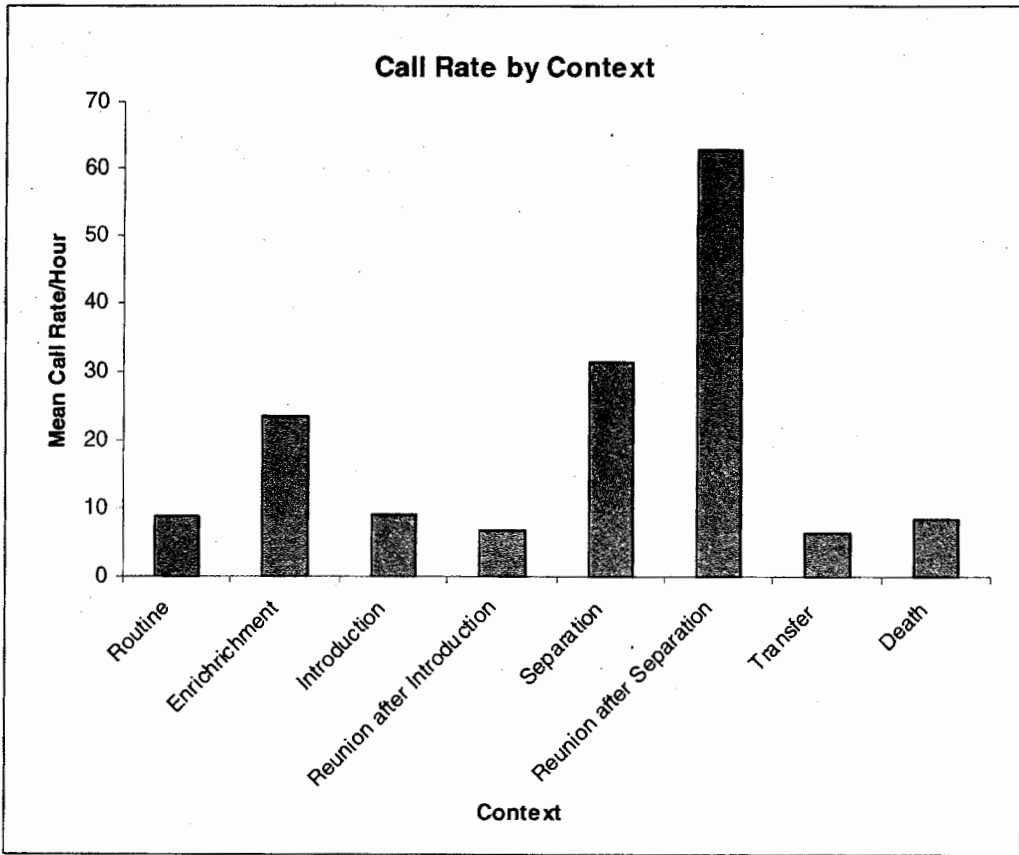


Figure 4: Mean call rate as a function of social context

## CHAPTER IV: ACOUSTIC PARAMETERS AND AUTOMATIC CLASSIFICATION OF CALLS

### *METHODS*

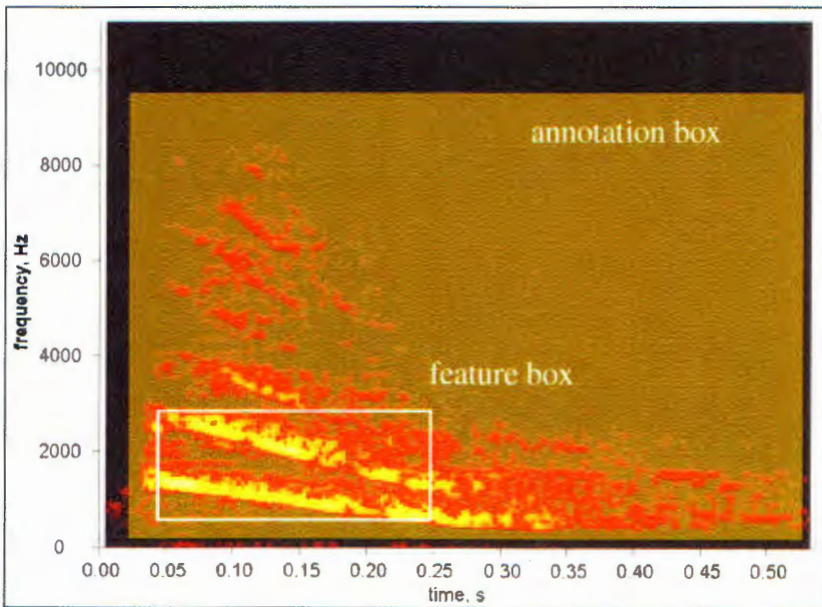
#### **Measuring acoustic parameters**

Praat scripts were run to make a separate sound file for each vocalizations.

Only sounds of basic call types (Bark, Roar, Rumble, Squeak, Squeals, Trumpet) and Blows with no overlapping sporadic sounds were selected for measurement. All calls were measured at the original sample rate of either 32 kHz or 44 kHz.

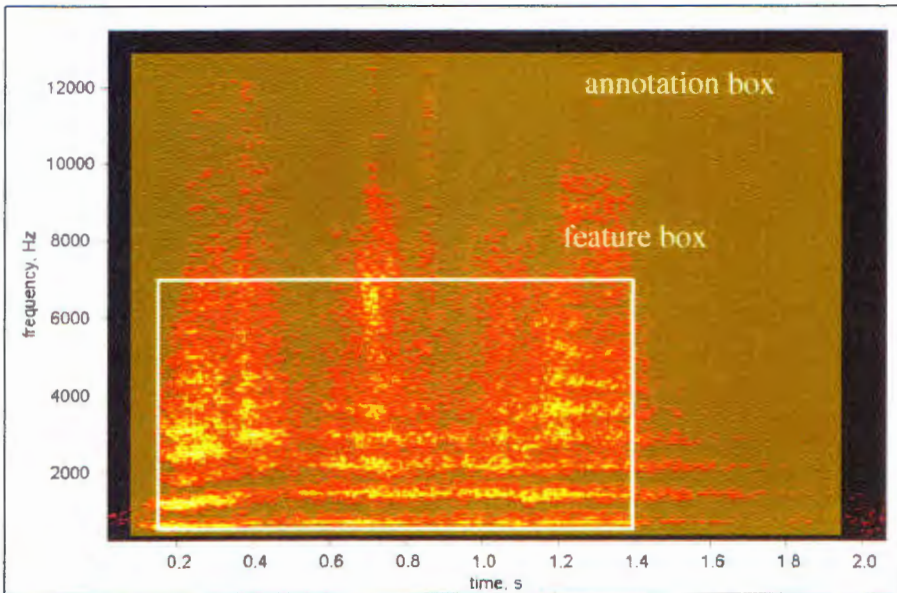
Osprey (v1.7) on MATLAB (v7.7, Mathworks, Inc.) was used to measure the acoustic parameters in order to characterize the signal structure of each call type. Osprey's measurement system was developed for characterizing the marine sounds in the Macaulay Library of Natural Sounds. Osprey measurements are based primarily on Fristrup and Watkin's (1993) AcouStat approach, with additional measurements and modifications that use estimators of central tendency and dispersion that are robust to outliers, namely the median, interquartile range, and quartile skewness. Measurements that use these estimators of central tendency have more consistent values at variable noise levels than measurements that rely on manual selection of signal extremes (Mellinger & Bradbury, 2007; Cortopassi, 2006). Signal extremes are sensitive to noise and outliers. In addition, manually measuring signal extremes requires assessment of signal onset and offset in both time and frequency, which can be affected by display settings and can be biased by researcher expectation and experience (Cortopassi, 2006).

Calls of the same type and sampling rate were merged together into a sequence of calls for quicker and more consistent measuring. Spacers were added between calls in the merged file for Rumbles and Squeals to help distinguish separate calls. The spectrogram parameters in Osprey were set to show separation of harmonics and good detail in both time and frequency, and were set the same for each call type (window type=Hamming, hop size= $\frac{1}{4}$ , zero padding=1x, frame size (digital samples per FFT) =512 to 2048, for a filter bandwidth of 63 to 349 Hz). An *annotation box* (or selection box) containing the entire sound in a region of time and frequency was drawn liberally to contain the entire sound plus some background noise, as shown in the two example calls in Figure 5 and Figure 6. Because only calls with no overlap of other sporadic sounds were used for measuring acoustic parameters, the annotation box included only the ambient noise and the focal sound.



**Figure 5: Osprey window of Squeak with annotation box and feature box.** X-axis is time. Y axis is frequency. Coloration is intensity. Annotation box encompasses the entire signal. Feature box encompasses 90% of the signal energy.





**Figure 6: Osprey window of Trumpet with annotation box and feature box.**

X-axis is time. Y axis is frequency. Coloration is intensity. Annotation box encompasses the entire signal. Feature box encompasses 90% of the signal energy.

Osprey creates a spectrogram of the annotation box and applies a de-noising algorithm to the entire spectrogram to remove general ambient noise (Mellinger & Bradbury, 2007). It then creates a *feature box* that encompasses the inner 90% of the signal energy in time and frequency, with the strongest 90% of the signal represented and the weakest 10% excluded. In the time domain (horizontal axis in the spectrogram), energy in each column is summed, with the sequence of sums forming a row vector known as the *time envelope*. The row vector is then ranked (sorted) highest to lowest in energy, and the cumulative sum beginning at the high end is computed until 90% of the total energy is reached. The earliest and latest time indices included in this 90% cumulative sum define the time bounds of the feature box. The inner 90% is the upper 90% after the values in the time envelope are sorted into ascending order. An analogous process happens in the frequency domain (vertical axis in the

spectrogram), with the sequence of row sums forming the *frequency envelope* (power spectrum), and the 90<sup>th</sup> percentile of the sorted energies being used to define the frequency bounds of the feature box.

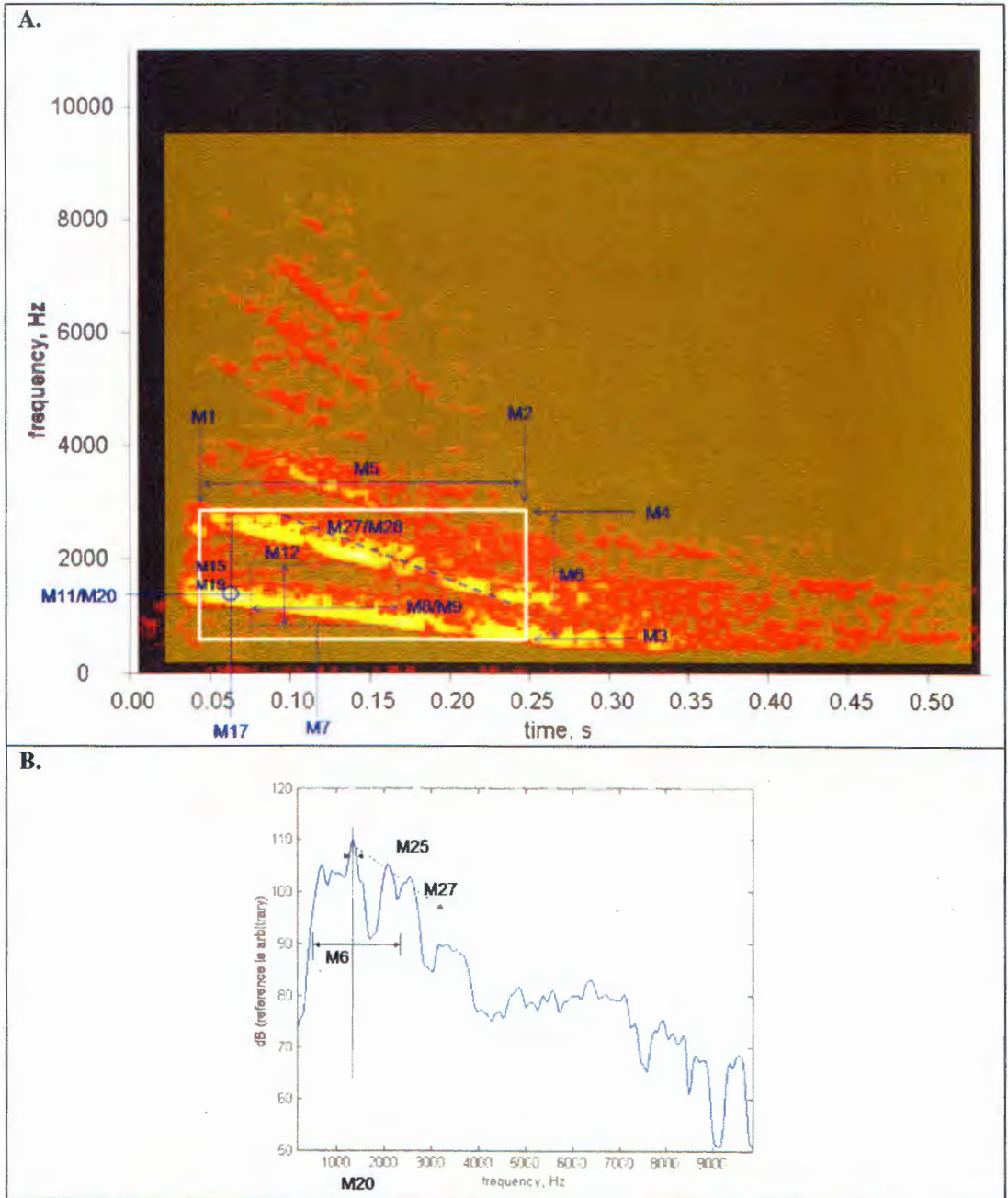
After the feature box is established, Osprey extracts 28 acoustic features within this feature box and calculates the signal to noise ratio within the annotation box.

Osprey weights measurements at each instant by the normalized intensity of the signal at that instant. This weighting means that louder parts, which are least affected by background noise, have the strongest influence on the measurement value (Mellinger & Bradbury, 2007).

For calls that have low frequency energy around 20 Hz (Rumbles, Barks, and Roars), the frequency of the fundamental was verified by measuring the harmonic spacing of the call in Praat. The annotation box was then drawn so the boundary matched the fundamental frequency rather than drawing it more liberally to include energy below the fundamental frequency. The reason for this was twofold: (1) one goal of this study was to determine the frequency range of elephant calls, so an accurate measure of the minimum frequency for low frequency calls was needed, and (2) drawing the annotation box below the fundamental frequency sometimes resulted in the inner 90% of the signal energy encompassing energy below the frequency that was verified as the fundamental frequency by harmonic spacing.

## Acoustic parameter definitions

The acoustic parameters measured by Osprey are illustrated in Figure 7 and described in Table 5. A complete description of the calculation of these parameters is given by Mellinger & Bradbury (2007).



**Figure 7: Acoustic parameters measured by Osprey that can be visualized**  
Panel A shows the spectrogram. Panel B shows the power spectrum.



**Table 3: Acoustic parameters measured by Osprey**

<i>Parameter</i>	<i>Units</i>	<i>Description</i>	<i>Type</i>
<b>M1-M6: Feature box</b>			
Start Time (M1)	sec	Lowest time index in bounds that encompass the inner 90% of the signal strength in the time envelope.	temporal
End Time (M2)	sec	Highest time index in bounds that encompass the inner 90% of the signal strength in the time envelope.	temporal
Lower Frequency (M3)	Hz	Lowest frequency index in bounds that encompass the inner 90% of the signal strength in the frequency envelope.	frequency
Upper Frequency (M4)	Hz	Highest frequency index in bounds that encompass the inner 90% of the signal strength in the frequency envelope.	frequency
Duration (M5)	sec	Width of feature box: M2-M1	temporal
Bandwidth (M6)	Hz	Height of feature box: M4-M3	frequency
<b>M7-M14: Central values and variation</b>			
<b>Uses measures that do not assume normality: median, quartile ranges, quartile skewness, concentration.</b>			
Median Time (M7)	sec	Time at which 50% cumulative signal energy is reached. (Measured relative to start of file, so M7 was calculated as $M7_{new}=M7-M1$ )	temporal
Temporal Interquartile Range (M8)	sec	Concentration of a call around the median time (M7) measured as the duration of the interquartile range of signal energy (Q3-Q1). Counts energy going forward and back from the median time (M7). Q3=median + 25% of signal energy Q1=median-25% of signal energy	temporal
Temporal Concentration (M9)	sec	Concentration of a call measured as the time span encompassing loudest 50% of time envelope values. Counts energy from the loudest parts down towards the smallest parts regardless of where the parts occur in time.	temporal
Temporal Asymmetry (M10)	none	Skewness of energy along time axis within interquartile range (-1.0 to 1.0)	temporal

<i>Parameter</i>	<i>Units</i>	<i>Description</i>	<i>Type</i>
Median Frequency (M11)	Hz	Frequency at with 50% cumulative signal energy is reached. More stable than extreme values of LowerFreq and UpperFreq in varying noise conditions.	frequency
Spectral Interquartile Range (M12)	Hz	Concentration of a call around the median frequency (M11) measured as frequency range of interquartile range of signal energy (Q3-Q1). Counts energy going forward and back from the median frequency (M11). Q3=median + 25% of signal energy Q1=median-25% of signal energy	frequency
Spectral Concentration (M13)	Hz	Concentration of a call measured as the frequency span encompassing loudest 50% of frequency envelope values. Counts energy from the loudest parts down towards the smallest parts regardless of where the parts occur in time.	frequency
Frequency Asymmetry (M14)	none	Skewness of energy along frequency axis within interquartile range (-1.0 to 1.0)	frequency
<b>M15-M20: Peak intensity</b>			
Time of Peak Cell Intensity (M15)	sec	Time of single loudest spectrogram cell. Time of the cell containing the peak intensity. (Measured relative to start of file, so M15 was calculated as $M15_{new}=M15-M1$ )	temporal
Relative Time of Peak Cell Intensity (M16)	%	Relative time of peak intensity ( $M15/M5$ )	temporal
Time of Peak Overall Intensity (M17)	sec	Largest value in time envelope, which is the largest vertical sum of the spectrogram over all frequencies. Time of the peak intensity in the trimmed time envelope. (Measured relative to start of file, so M7 was calculated as $M17_{new}=M17-M1$ )	temporal
Relative Time of Peak Overall Intensity (M18)	%	Relative time of peak intensity ( $M17/M5$ )	temporal
Frequency of Peak Cell Intensity (M19)	Hz	Frequency of cell containing the peak intensity.	frequency
Frequency of Peak Overall Intensity (M20)	Hz	Frequency of peak intensity in the trimmed frequency envelope.	frequency

<i>Parameter</i>	<i>Units</i>	<i>Description</i>	<i>Type</i>
<b>M21-M24: Amplitude and frequency modulation (variation of amplitude and frequency over time)</b>			
AM Rate (M21)	Hz	Dominant rate of amplitude modulation. Frequency of the maximum rate in the power spectrum of the trimmed time envelope.	amplitude
AM Rate Variation (M22)	Hz	Variability of amplitude modulation measured as the width of peak at M21-6 dB. Values are discretized because at 6dB down from the peak, the widths may be a only a few bins wide so the values are integer multiples of the bin width.	amplitude
FM Rate (M23)	Hz	Dominant rate of frequency modulation. Frequency of the maximum rate in the power spectrum of the trimmed frequency envelope.	frequency
FM Rate Variation (M24)	Hz	Variability of frequency modulation measured as the width of peak at M23-6 dB. (How much the rate of change varies, may be related to inflections and steepness of upsweeps and downsweeps)	frequency
<b>M25-M28: Fine features of harmonic structure, shifts in periodicity, direction of frequency change, rate of change in frequency</b>			
Cepstrum Peak Width (M25)	Hz	Harmonic structure Average width of peaks (harmonics) in power spectrum. Peak width is measured at 6 dB down from maximum value. At 6 dB down from the peak, the widths may be a only a few bins wide (like M22 and M24), but M25 is an average of integers so the values are not discretized. Narrow peaks means narrowband/tonal harmonics.	structure
Overall Entropy (M26)	Hz	Entropy, shifts in periodicity Distribution of energy across frequency blocks in a given time block. Shift from periodicity and linearity to chaos. Change in noisiness v. tonality.	structure
Upsweep Mean (M27)	Hz	Direction of frequency change Measures how much the frequency increases. Average change in median frequency between successive time blocks, weighted by total energy in the block. Inflection points with rising and falling frequencies throughout call result in a low M28 (closer to 0) compared to a consistent directional change. Measure is weighted to emphasize contribution of louder signal components. M27<0 means frequency is decreasing.	frequency

<i>Parameter</i>	<i>Units</i>	<i>Description</i>	<i>Type</i>
Upsweep Fraction (M28)	%	Rate of directional frequency change Counts the number of times the frequency content increases. Fraction of time in which the median frequency in one block is greater than in the preceding block, weighted by total energy in the block. Indicates how much of the call has a directional change in the frequency. Inflection points with rising and falling frequencies throughout call result in a high M28, just as a consistent directional change. Measure is weighted to emphasize contribution of louder signal components. M28 always positive.	frequency
<b>M29: Signal strength</b>			
Signal-to-Noise Ratio (M29)	dB	Signal to noise ratio within the annotation box. Ratio of the signal power (loudest cell) to the noise power (power of cell at 25 <sup>th</sup> percentile). Cells are ranked low to high and the cell at the 25 <sup>th</sup> percentile represents noise. (25 <sup>th</sup> percentile is used because the animal call likely takes up less than 75% of the total spectrogram cells.) Measurement assumes that the within the annotation box at least 25% of the cells are without a focal signal.	amplitude

### **Dataset for measuring acoustic parameters**

The original dataset contained 2791 calls representing all call types, including combination calls. Acoustic parameters were measured for basic call types (Bark, Roar, Rumble, Squeak, Squeals, Trumpet) and Blows having no overlapping sporadic sounds and no echo (N=1011). Calls that met the following criteria were removed from this dataset for statistical analysis: duplicate calls using different recording systems (N=83) and calls with a low signal-to-noise ratio (SNR < 10 dB) (N=1). The final dataset used for statistical analysis is shown in Table 4. It is important to note that the number of samples per call type was not balanced, so the statistical power is

low. Also, Blow sounds were not scored for every occurrence, so the sample size was very low relative to the number of occurrence.

**Table 4: Calls in dataset for statistical analysis (N=927)**

Call type	Oregon Zoo (N)	Thailand (N)	Total (N)
Blow	73	5	78
Bark	55	8	63
Roar	98	10	108
Rumble	38	21	59
Squeak	193	5	198
Squeal (includes Squeal, Squeal train, and Squeal long,)	211	31	242
Trumpet	123	56	179
TOTAL	791	136	927

### **Process for statistical analysis for call classification**

The goal of the statistical analysis was to determine if the data support automatic classification of calls that were previously categorized using perceptual aural cues and visual inspection of spectrograms, and to determine which acoustic parameters differentiate the call types. The process for the analysis was as follows: 1) screen data to investigate distributions and variability, 2) standardize data to account for different units of measure, 3) determine parameters to use for statistical analysis, 4) remove outliers, 5) determine parametric or non-parametric statistical analyses, and 6) run statistical analysis for automatic classification. All statistics were run using R (v2.8.1).

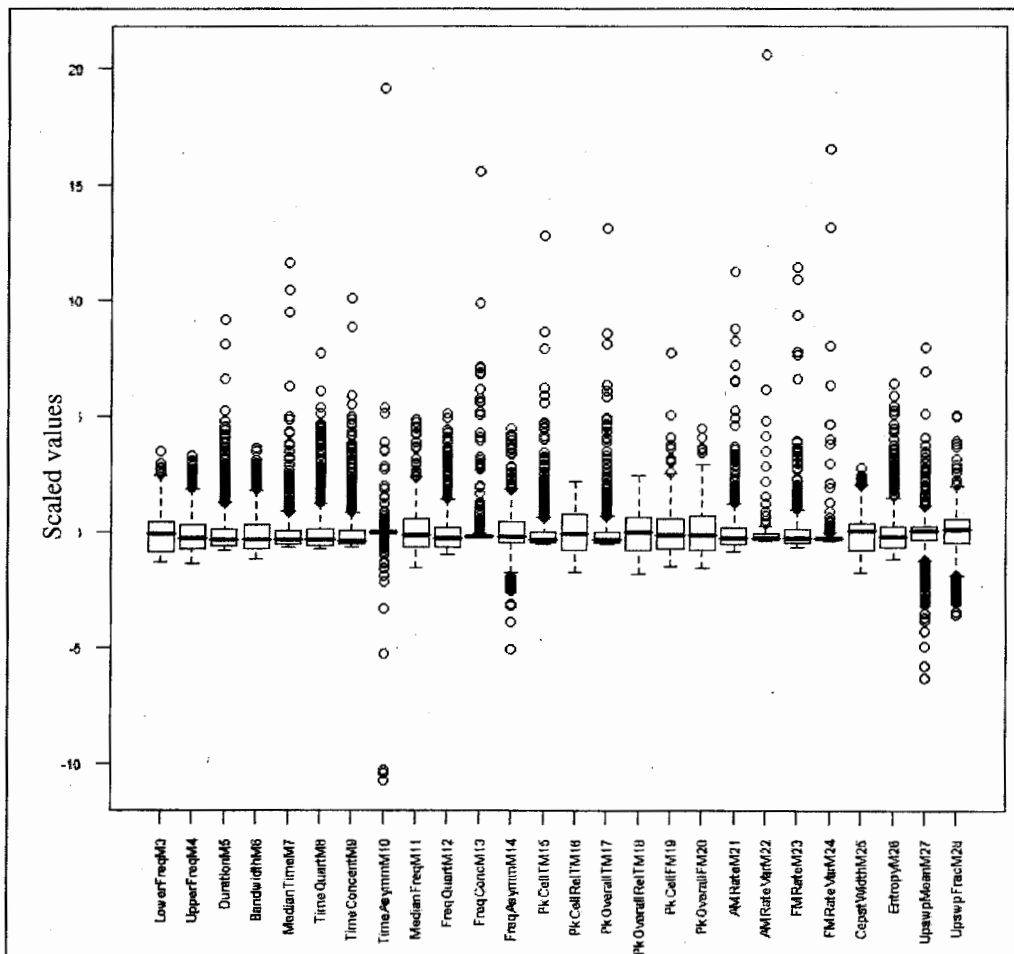
### **Screening data for statistical analysis**

Histograms and boxplots were created to determine the distribution and variability among call types, to do a preliminary assessment of variables that

differentiate call types, and to learn more about how these variables reflect perceptual aural cues.

### Standardizing the data for statistical analysis

The acoustic parameters are measured in different units, so the data must be standardized, or scaled, for the variables to be dimensionally homogeneous. A boxplot of all variables was created for the scaled dataset, as shown in Figure 8.



**Figure 8: Boxplot of scaled variables across all call types (blow, bark, roar, rumble, squeak, squeal, trumpet)**  
Acoustic parameters are on the x-axis. Scaled (normalized) values of the parameters are on the y-axis.

## **Removing parameters from statistical analysis**

Some parameters were removed from the statistical analysis because they were highly correlated with other parameters for this dataset (6 parameters) or were absolute measures that could be represented by relative measures (4 parameters).

A correlation matrix was created using the scaled dataset of the basic call types (Bark, Roar, Rumble, Squeak, Squeal, Trumpet) and Blow to determine potential candidates for removal from further analyses. The correlation matrix measures the strength and direction of the linear relationship between pair wise combinations of variables (given by the Pearson correlation coefficient). Pairs of acoustic parameters having high correlation coefficients were considered candidates for removal. Because a correlation measures only the strength of linear relationships, it is possible that additional pair wise combinations had a strong relationship, but the relationship was not linear.

Correlations were greater than 0.5 for 14 pairs of variables. Although high correlations could indicate redundancy and therefore suggest candidates for removal from further analyses, not all candidates were removed because some relationships may be biologically relevant for investigating acoustic communication of this species as compared to others. For example, Overall Entropy (M26) was highly positively correlated with three variables related to frequency range, Upper Frequency (M4), Bandwidth (M6), and Median Frequency (M11), so M26 was a candidate for removal. However, the relationship of entropy to frequency range may provide insight into sound production and perception. Frequency of Peak Overall Intensity (M20) was highly correlated with Low Frequency (M3), so M20 was a candidate for removal.

However, this relationship confirms that the peak intensity occurs at or close to the fundamental frequency of the call. Other high correlations were related to central values and variation and their relationship to frequency and duration. These correlations may indicate greater stereotypy in elephant calls compared to calls of other species, so the relationship between these variables may be of biological interest. Further analysis may be needed to determine measurements that are most important given the study subjects and recording environment.

A total of 18 acoustic parameters were used in the statistical analyses, which are shown in Table 5 with notation for each parameter that appears in the data output.

**Table 5: Acoustic parameters that were used in statistical analysis**

<i>Parameter</i>	<i>Notation in Output</i>	<i>Description</i>
Lower Frequency (Hz) (M3)	LowerFreqM3	Lowest frequency index in bounds that encompass the inner 90% of the signal strength in the frequency envelope.
Upper Frequency (Hz) (M4)	UpperFreqM4	Highest frequency index in bounds that encompass the inner 90% of the signal strength in the frequency envelope.
Duration (sec) (M5)	DurationM5	Width of feature box: M2-M1
Bandwidth (Hz) (M6)	BandwidthM6	Height of feature box: M4-M3
Median Time (sec) (M7)	MedianTimeM7	Time at which 50% cumulative signal energy is reached. (Measured relative to start of file, so M7 was calculated as $M7_{new}=M7-M1$ )
Temporal Concentration (sec) (M9)	TimeConcentM9	Concentration of a call is measured as the time span encompassing loudest 50% of time envelope values. Counts energy from the loudest parts down towards the smallest parts regardless of where the parts occur in time.
Temporal Asymmetry (M10)	TimeAsymmM10	Skewness of energy along time axis within interquartile range (-1.0 to 1.0)

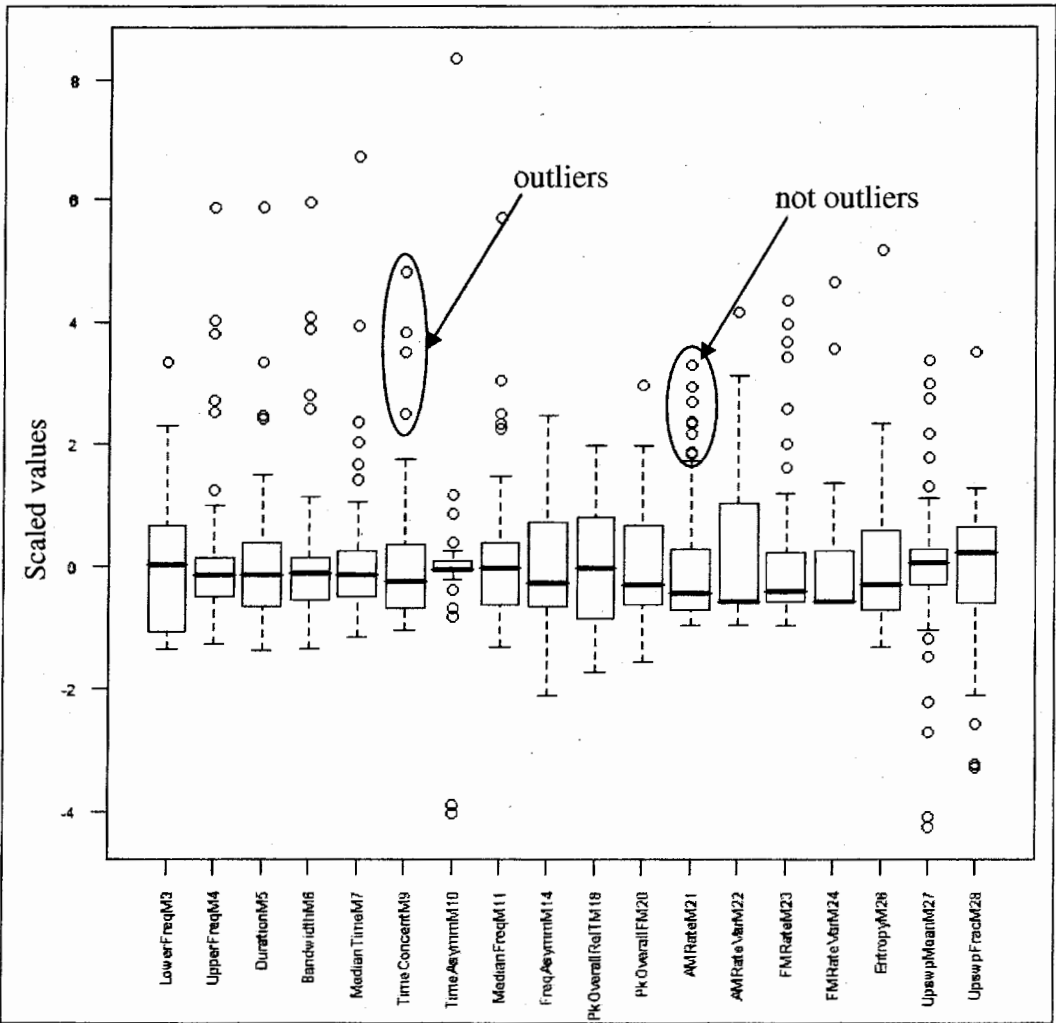


<i>Parameter</i>	<i>Notation in Output</i>	<i>Description</i>
Median Frequency (Hz) (M11)	MedianFreqM11	Frequency at with 50% cumulative signal energy is reached. More stable than extreme values of LowerFreq and UpperFreq in varying noise conditions.
Frequency Asymmetry (M14)	FreqAsymmM14	Skewness of energy along frequency axis within interquartile range (-1.0 to 1.0)
Relative Time of Peak Overall Intensity (%) (M18)	PkOverallRelTM18	Relative time of peak intensity (M17/M5)
Frequency of Peak Overall Intensity (Hz) (M20)	PkOverallFM20	Frequency of peak intensity in the trimmed frequency envelope.
AM Rate (Hz) (M21)	AMRateM21	Dominant rate of amplitude modulation. Frequency of the maximum rate in the power spectrum of the trimmed time envelope.
AM Rate Variation (Hz) (M22)	AMRateVarM22	Variability of amplitude modulation measured as the width of peak at M21-6 dB. Values are discretized because at 6dB down from the peak, the widths may be a only a few bins wide so the values are integer multiples of the bin width.
FM Rate (Hz) (M23)	FMRateM23	Dominant rate of frequency modulation. Frequency of the maximum rate in the power spectrum of the trimmed frequency envelope.
FM Rate Variation (Hz) (M24)	FMRateVarM24	Variability of frequency modulation measured as the width of peak at M23-6 dB. (How much the rate of change varies, may be related to inflections and steepness of upsweeps and downsweeps)
Overall Entropy (Hz) (M26)	EntropyM26	Entropy, shifts in periodicity Distribution of energy across frequency blocks in a given time block. Shift from periodicity and linearity to chaos. Change in noisiness v. tonality.
Upsweep Mean (Hz) (M27)	UpswpMeanM27	Direction of frequency change Measures how much the frequency increases. Average change in median frequency between successive time blocks, weighted by total energy in the block. Inflection points with rising and falling frequencies throughout call result in a low M28 (closer to 0) compared to a consistent directional change. Measure is weighted to emphasize contribution of louder signal components. M27<0 means frequency is decreasing.

<i>Parameter</i>	<i>Notation in Output</i>	<i>Description</i>
Upsweep Fraction (Hz) (M28)	UpswpFracM28	Rate of directional frequency change Counts the number of times the frequency content increases. Fraction of time in which the median frequency in one block is greater than in the preceding block, weighted by total energy in the block. Indicates how much of the call has a directional change in the frequency. Inflection points with rising and falling frequencies throughout call result in a high M28, just as a consistent directional change. Measure is weighted to emphasize contribution of louder signal components. M28 always positive.

## Removing outliers

Boxplots of all variables in the scaled dataset were created for each call type to determine potential outliers within each call type. The criteria for determining outliers was very conservative in order to preserve data that may have biological relevance. Given that the gradients of variability may have biological relevance, data points were considered outliers only if they deviated from the overall pattern of distribution and did not appear to belong to a long tail (gradient), as shown in Figure 9. Calls were considered outliers only if they were multivariate outliers. Because variance across two variables could be a result of covariance alone, calls were considered outliers and removed from the dataset only if they were outliers across at least three variables. Because calls recorded in Thailand could represent variability in call structure not found in the Oregon Zoo call repertoire, and the Oregon Zoo data is better represented in the dataset, outlier Thailand calls were removed only if they sounded like they did not meet the definition of the calls as established by aural cues.



**Figure 9: Scaled variables for Roar showing outliers and gradients**

Small circles in the plot indicate points that are outside of the interquartile range. Ellipses show examples of data points considered as outliers and data points not considered outliers.

Out of 927 calls, only 19 calls (Blow N=6, Bark N=4, Roar N=3, Rumble N=2, Squeak N=1, Squeal N=3, Trumpet N=0) met the conservative criteria for outliers.

Table 6 shows the final dataset with outliers removed (N=908).

**Table 6: Calls in dataset (N=908) after outliers removed**

Call type	Oregon Zoo (N)	Thailand (N)	Total (N)
Blow	68	4	72
Bark	51	8	59
Roar	97	8	105
Rumble	36	21	57
Squeak	193	4	198
Squeal (includes Squeal, Squeal train, and Squeal long,)	209	30	239
Trumpet	123	56	179
TOTAL	777	131	908

### **Determining parametric or non-parametric statistical analysis**

After scaling and removal of outliers, normal probability plots were created for the scaled dataset (N=908) to determine how well the normal distribution describes the data. The normal probability plot is a quantile-quantile (Q-Q) plot against the standard normal distribution. The observed data is ranked and the quantile is calculated. If the observed data set matches the theoretical distribution, the shape of the plot will be a straight line where  $y = x$ .

Based on inspection of the Q-Q plots, approximately half of the variables were normally distributed and only slightly skewed, and half were either very skewed or not normal. Therefore, this dataset does not meet the assumptions of normality, linearity, and constant variance. By not meeting these assumptions, only non-parametric statistics can be used.

## **Running statistical analyses for automatic classification**

A classification tree was run with the original data (non-standardized) as a preliminary analysis for determining the most important predictors for classification into the basic call types (Bark, Roar, Rumble, Squeak, Squeals, Trumpet), and to explore its potential as a predictive model for classifying new calls into the pre-defined call types.

A Principal Component Analysis (PCA) of the scaled dataset was run to determine the parameters that explain most of the variance among call types. The main objective of PCA is to reduce the dimensionality of the dataset. Ideally, PCA should be multinormal and reasonably unskewed. Although the data did not meet these assumptions, the clustering of data along principal component axes can still provide insight into the variables that group and separate call types.

Global and pair-wise Analysis of Similarity (ANOSIM) tests of the scaled dataset were run to determine if there were significant differences in the acoustic parameters among the pre-defined call types and between pairs of call types. If two call types are significantly different in their parameter values, then the dissimilarities between the call types will be greater than those within each call type. In order to evaluate the dissimilarity within and between call types as a measure of true distance between parameter values, the Ward's linkage method was used with the Euclidian distance. The number of iterations for assessing significance was 1000. A Bonferroni correction was calculated to reduce the family-wise Type I error from 15% to 4.8%.

Other analyses considered were a Discriminant Function Analysis (DFA) and unsupervised clustering. DFA is better than PCA for discriminating existing groups

and classifying into pre-defined categories; however, even the scaled data did not meet the assumption of multinormality required by DFA. A hierarchical agglomerative cluster analysis of the scaled dataset was run on the Oregon Zoo data using R.

Although there appeared to be at least a weak group structure in support of the pre-defined call types, the output plot was illegible given the number of data points so it was not possible to determine other potential call groupings from inspection.

Unsupervised clustering should be considered for future comparison of call classification methods.

The Blow sound was included in the distributions, but not the classification tree, PCA, or ANOSIM tests. This non-vocal sound is made frequently by both males and females and in most situations, so characterizing this sound by its acoustic structure may be useful for acoustic monitoring. However, preliminary runs of the classification tree and PCA showed that it complicated the classification and grouping of the vocal sounds, so it was handled separately from the basic call types.

### **Comparing low frequency calls to background noise**

Although low frequency vocalizations have the potential for communication in forested environments and over long distances, they serve a communicative function only if they can be detected by a receiver. Power spectra of low-frequency Rumble calls and background noise were compared to investigate bandwidth and intensity of Rumbles in relation to spectral characteristics and intensity of the background noise.

Only data collected with equipment capable of recording below 20 Hz were used. Paired Rumble and noise samples were selected from the same recording to

ensure that the recording levels were equal. Noise segments were selected as close as possible to the Rumble to take into account that a caller may adjust the intensity of sound production depending on background noise. The noise segments were selected with a duration approximately equal to the rumble. A frequency range of 14 Hz to 1.2 kHz was used to compare bandpass characteristics because rumbles have a range of 14 Hz to 1.2 kHz for the inner 90% of the call energy. Thailand recordings were separated by site because one site is rural (Elephant Nature Park) and the other is urban (Royal Elephant Kraal).

Osprey uses an arbitrary reference to generate power spectra, where a sample value of 1 is effectively 0 dB. Hence, comparisons of signal and noise can be made only if the recording levels are equal so that the power is referenced to the same value. With this reference to the same value, the signal to noise ratio at any frequency value is the difference of the signal power (dB) and noise power (dB). To make these comparison, the signal and noise were not separated within each sound segment; rather, the spectrums for the paired Rumble and noise segments were compared visually to estimate the difference between the signal power and the noise power at various frequency values.

The sample size used in these analysis was small (Oregon Zoo N=3, Elephant Nature Park N=4, Royal Elephant Kraal N=3), so this does not constitute a complete noise analysis, but it does offer some insight into the potential of Rumble calls to be buried in noise in various recording environments. Future analyses include a more

quantitative analysis of the signal to noise ratio at various frequency values for an increased number of call and noise samples.

## **RESULTS**

### **Call types**

Six basic vocalization types were identified as structurally different, and were labeled Bark, Roar, Rumble, Squeak, Squeal, Trumpet. An additional non-vocal sound, Blow, was also found.

### **Frequency range, bandwidth, and duration of acoustic repertoire**

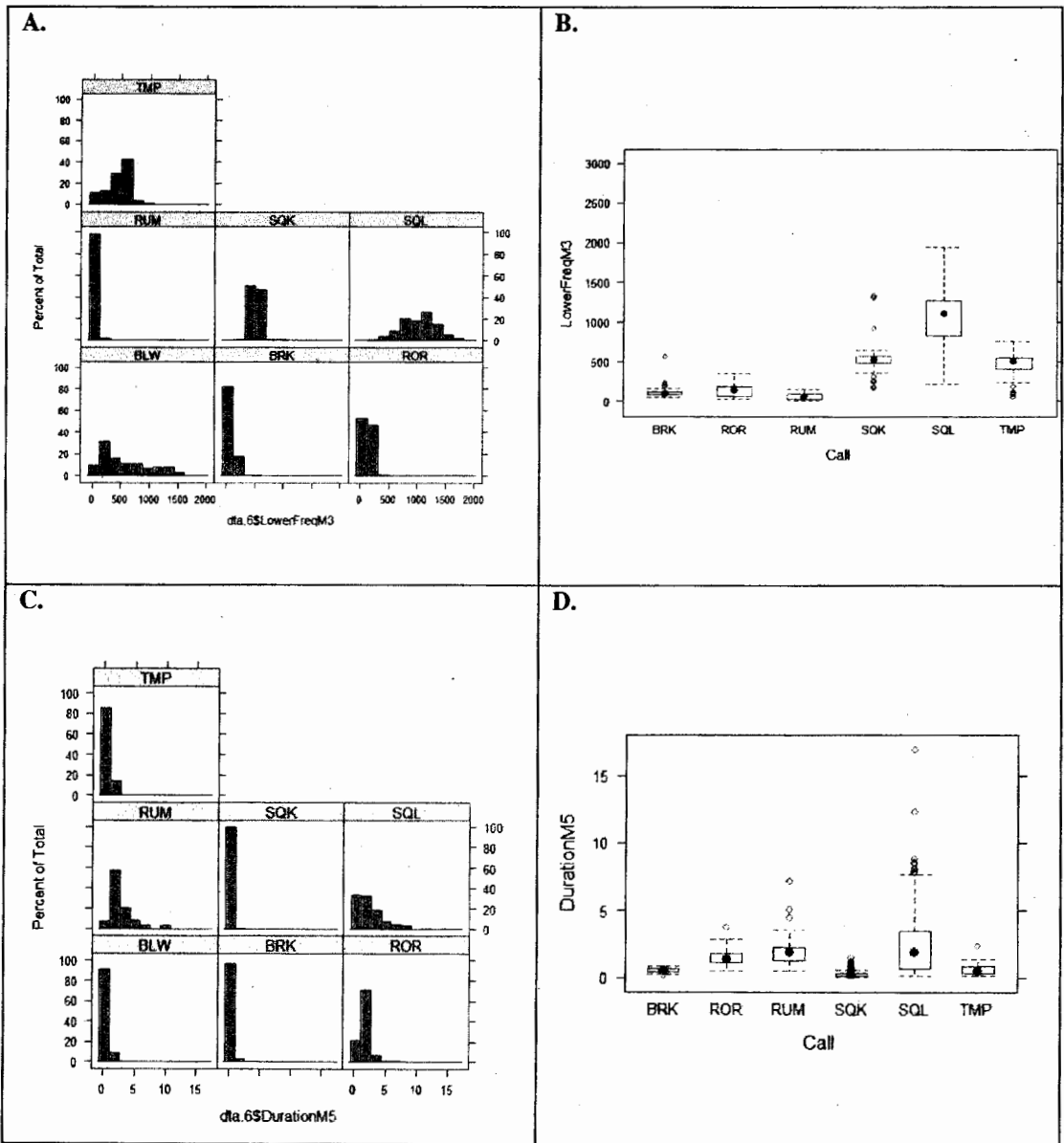
The frequency range of the basic call types (Bark, Roar, Rumble, Squeak, Squeal, Trumpet) and Blow encompassing all of the signal energy was 14 Hz to 18 kHz. The frequency range encompassing only 90% of the energy was 14 Hz to 9 kHz. The bandwidth encompassing 90% of the signal energy for individual calls ranged from 54 Hz to 9 kHz (median 1680 Hz). The duration encompassing 90% of the signal energy varied from 0.1 sec to 14 seconds (median 0.7 sec).

### **Variability of acoustic parameters among call types**

Only a subset of the graphical results are included here. All plots are provided in the appendices. By comparing the histograms and boxplots, it appears that the distribution of each variable differs among call types, as shown in Figure 10. From the distributions of commonly used measurements (e.g., fundamental frequency and duration), one can see how the acoustic parameters reflect aural perception. For example, the fundamental frequency, as measured by Lower Frequency (M3), is



consistent or stereotypic for Bark (BRK), Roar (ROR), Rumble (RUM) and Squeak (SQK), but varies more widely for Trumpet (TMP) and even more so for Squeal (SQL). Blow shows the greatest variance, but this may be because it is a non-vocal sound that is mostly broadband noise. Consistent with perception by aural cues, Bark, Roar, and Rumble are lower in frequency than Squeak, Squeal, and Trumpet. Most of the energy in Rumbles is low, whereas Roars have energy that extends higher. The Duration (M5) is stereotypic for Trumpet, Squeak, Blow, and Bark, but is highly variable for Rumble and Squeal. Squeal include the longest calls, followed by Rumble then Roar.



**Figure 10: Example histograms and boxplots of acoustic parameters for each call type show distribution and variability.**

TMP=Trumpet, RUM=Rumble, SQK=Squeak, SQL=Squeal, BLW=Blow, BRK=Bark, ROR=Roar. Panel A and B show the Low Frequency (M3) parameter values or each call type. Panel C and D show the Duration (M5) parameter values for each call type.

## Classification tree

A classification tree was run on the non-standardized data (N=836), with outliers removed and excluding Blows. The tree served as a preliminary analysis for determining the most important predictors for classification, and may have potential as a predictive model for classifying new calls into the pre-defined call types (Bark, Roar, Rumble, Squeak, Squeal, Trumpet). The final classification tree model is shown in Figure 10. Tree pruning was used as cross-validation for determining the final model. The goal of cross-validation was to evaluate alternative models that reduce the number of branches, but when the number of branches was reduced, two call types (Bark and Rumble) were left out and the misclassification rate doubled. The final model explained 78.4% of the variation among call types using decision rules with only six variables: Lower Frequency (M3), Upper Frequency (M4), Duration (M5), Frequency of Peak Overall Intensity (M20), FM Rate Variation (M24), and Upsweep Fraction (M28).

The overall misclassification rate was 12.2%, which means successful classification of 87.8% of the calls into the predicted call type. Table 7 and Table 8 show the confusion matrix and the classification rate for each call type. The Squeak had the highest successful classification rate (95.9%), followed by Squeal (91.2%), Trumpet (91.1%), Rumble (78.9%), Roar (73.3%), then Bark (71.1%). Rumbles were most often misclassified as Roars. Roars were most often misclassified as Rumbles. Barks were most often misclassified as Trumpets. The sample sizes of Bark and

Rumble were low, which may impact the decision rules and thus successful classification of these call types.

The tree was most successful in classifying the group of higher frequency calls (Squeak, Squeal, and Trumpet), but classifying Squeaks and Squeals required multiple decision rules. Squeals were best differentiated by having a Lower Frequency (M3) above 675 Hz, but for those Squeal calls that did not meet this criteria, it appears that Upper Frequency (M4), Duration (M5), and FM Rate Variation (M24) differentiated this call type. The Squeal call type is highly variable, and it appears from the decision rules that misclassifications are primarily with the short Squeal call, which are often low intensity compared to Squeal trains and long Squeals. Squeaks were best differentiated by a low value for FM Rate Variation (M24), which means the frequency modulation rate was relatively consistent throughout the call. For those Squeaks that did not meet this criteria, it appears that Upper Frequency (M4), Upsweep Fraction (M28), Duration, and Lower Frequency (M3) differentiated this call type.

The group of lower frequency calls (Bark, Roar, and Rumble) appears to be primarily differentiated by low values of Lower Frequency (M3) and Upper Frequency (M4), then by Duration (M5). Among these call types, the Bark is separated from the Roar and Rumble by its short duration. The Roar and Rumble are separated by the Frequency of Peak Overall Intensity (M20), with the Rumble being below 170 Hz. Given that Roars and Rumbles were misclassified most often as each other, this single parameter may not be sufficient for separating these call types.

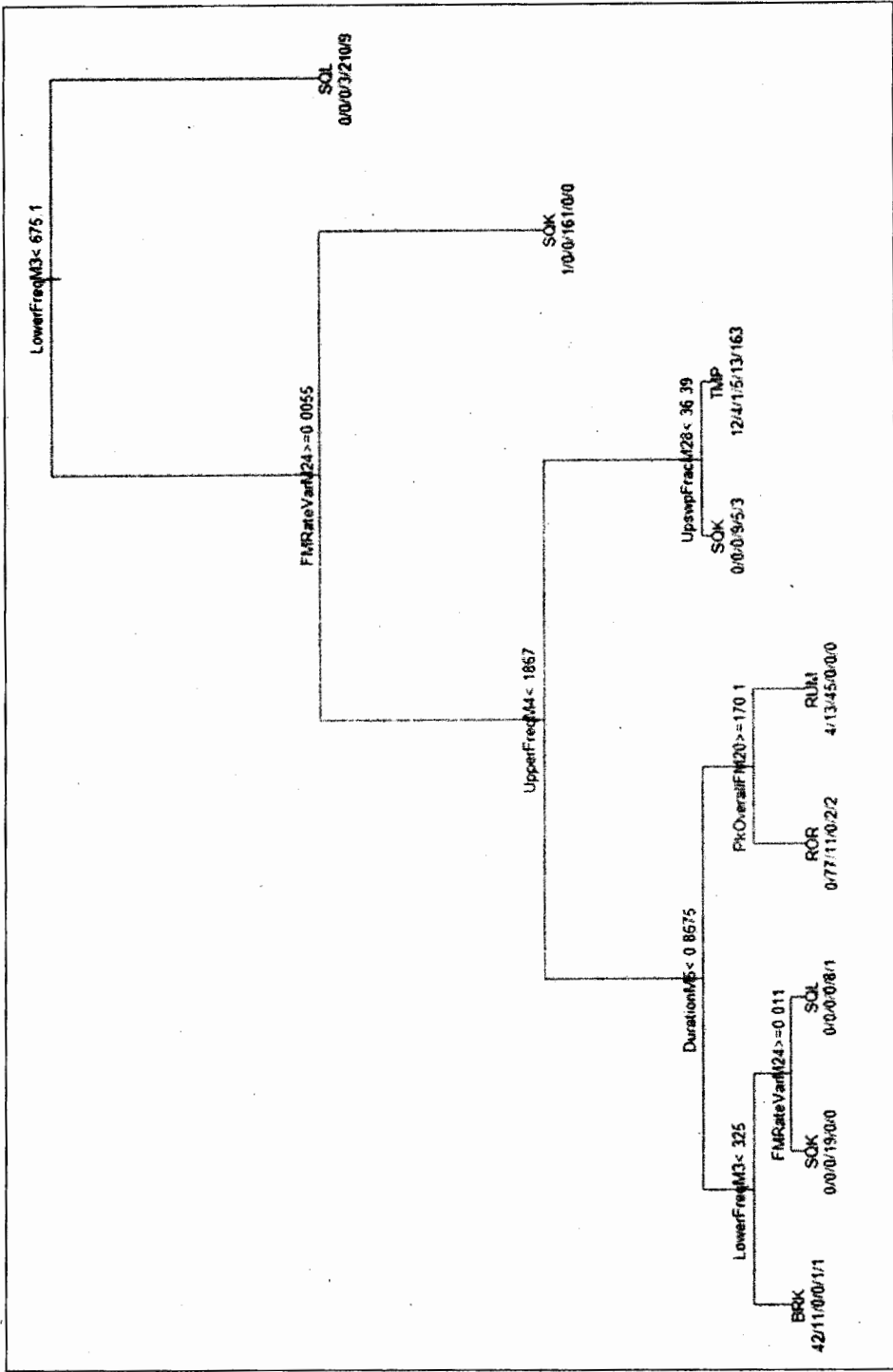


Figure 11: Final classification tree model.  
 (TMP=Trumpet, RUM=Rumble, SQK=Squeak, SQL=Squeal, BLW=Blow, BRK=Bark, ROR=Roar)

**Table 7: Confusion matrix for tree**

Diagonal cells (in grey) show correct classification into the predicted class. Off-diagonal cells show misclassifications. Columns represent the distribution of calls into each leaf (call type). Rows represent the distribution of each call type (collection of call types).

	Bark	Roar	Rumble	Squeak	Squeal	Trumpet	
Bark	42	11	0	0	1	1	Bark class contains 42 barks, 11 roars, 1 squeal, 1 trumpet
Roar	0	77	11	0	2	2	
Rumble	4	13	45	0	0	0	
Squeak	1	0	0	189	5	3	
Squeal	0	0	0	3	218	10	
Trumpet	12	4	1	5	13	163	
	59	105	57	197	239	179	# calls of each call type

↑ 59 bark calls recognized by tree as bark (42), rumble (4), squeak (1) trumpet (12)

**Table 8: Classification rate (%) for tree**

Diagonal cells (in grey) show correct classification into the predicted class. Off-diagonal cells show misclassifications

	Bark	Roar	Rumble	Squeak	Squeal	Trumpet	
Bark	71.1	10.5	0	0	0.4	0.6	Bark calls classified correctly at 71.1% with a 28.9% misclassification rate
Roar	0	73.3	19.3	0	0.8	1.2	
Rumble	6.7	12.4	78.9	0	0	0	
Squeak	1.7	0	0	95.9	2.1	1.7	
Squeal	0	0	0	1.3	91.2	5.6	
Trumpet	20.3	3.8	1.8	2.1	5.4	91.1	

Parameters measured by Osprey are numerous, and one could limit the parameters used in the model to those commonly measured by spectrographic analysis tools to design a tree that could be used widely by researchers of Asian elephant calls.

### Principal Component Analysis

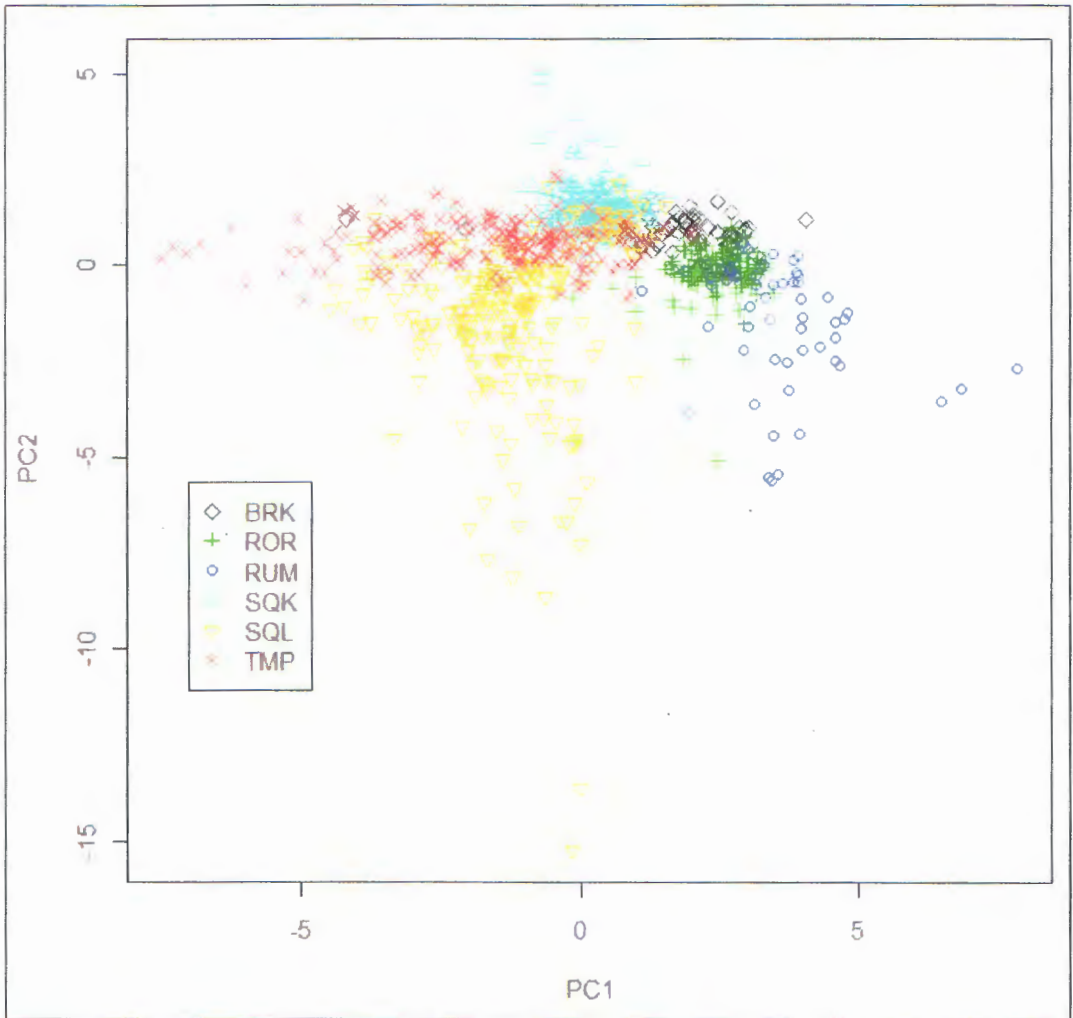
A Principal Component Analysis (PCA) of the scaled dataset (N=836) with outliers removed and excluding Blows was run to determine the parameters that account for the variance within and among call types. In accordance with broken.stick validation, which determines whether the observed pattern is significantly different from a random pattern, only PC1 and PC2 should be kept. PC1 and PC2 explained

only 42.8% of the variance, with PC1 explaining 24.6% of the variance and PC2 explaining 18.2%.

The PCA plot in Figure 12 shows some support for the grouping of call types, but with mixing of call types. The relative contribution of each variable to the principle components, or loading of principal components, highlights a subset of the variables as being important for separating call types and grouping calls within each type. The sign of the numerical loading value indicates whether the value of that parameters increases or decreases along the principal component axes. The loading is expressed in the eigenvectors of PC1 and PC2 below:

$$\text{PC1} = (-0.29)\text{LowerFreqM3} + (-0.42)\text{UpperFreqM4} + (-0.38)\text{BandwidthM6} + (-0.44)\text{MedianFreqM11} + (-0.38)\text{PkOverallFM20} + (0.23)\text{AMRateVarM22} + (0.21)\text{FMRateVarM24} + (-0.38)\text{EntropyM26}$$

$$\text{PC2} = (-0.14)\text{LowerFreqM3} + (-0.52)\text{DurationM5} + (-0.51)\text{MedianTimeM7} + (-0.51)\text{TimeConcentM9} + (-0.10)\text{PkOverallFM20} + (0.24)\text{AMRateM21} + (0.24)\text{FMRateM23} + (-0.19)\text{UpswpFracM28}$$



**Figure 12: Principal Component Analysis plot for basic call types**

TMP=Trumpet. RUM=Rumble. SQK=Squeak. SQL=Squeal. BLW=Blow. BRK=Bark. ROR=Roar  
 PC1 is represented primarily by spectral parameters, so clustering along PC1 results primarily from variability in spectral structure. PC2 is represented primarily by temporal parameters, so clustering along PC1 results primarily from variability in temporal patterns.

Along PC1 axis, the variability is explained by AM Rate Variation (M22), Overall Entropy (M26), and frequency parameters, namely Lower Frequency (M3), Upper Frequency (M4), Bandwidth (M6), Median Frequency (M11), Frequency of Peak Overall Intensity (M20), and FM Rate Variation (M24). The loading of AM Rate Variation (M22) and FM Rate Variation (M24) is positive, so the value of these



parameters increase along the PC1 axis. All other variables decrease along the PC1 axis. Starting at the most positive end of the PC1 axis, Rumbles have the lowest values for Lower Frequency (M3), Upper Frequency (M4), Bandwidth (M6), Median Frequency (M11), Frequency of Peak Overall Intensity (M20), and Overall Entropy (M26), and the highest values for FM Rate Variation (M24) and AM Rate Variation (M22). Moving from right to left along the PC1 axis are Rumbles, Roars, Barks, Squeaks, then Trumpets and Squeals with the highest values for Lower Frequency (M3), Upper Frequency (M4), Bandwidth (M6), Median Frequency (M11), Frequency of Peak Overall Intensity (M20), and Overall Entropy (M26), and the lowest values for Frequency modulation (M24) and Amplitude modulation (M22).

Squeaks show the least within-call variation along the PC1 axis, so this call type is the most stereotypic in its spectral structure. Trumpets show the greatest degree of variation along the PC1 axis, so this call type is highly variable in its spectral structure. These results are consistent with perceptual aural cues of lower and higher frequency calls, and of stereotypy in the Squeak call. It is difficult to perceive rate of change in modulation rates, so the results here provide differentiation that would otherwise be missed. Frequency jumps were perceived in the Squeal and changes in tonality and noisiness were perceived in the Trumpet, so the results here for the variability in Entropy are consistent with aural cues for these call types.

Along the PC2 axis, the variability is explained primarily by AM Rate (M21), FM Rate (M23), and temporal parameters, namely Duration (M5), Median Time (M7), Temporal Concentration (M9). There is some contribution by Lower Frequency (M3),

Frequency of Peak Overall Intensity (M20), and Upsweep Fraction (M28), but the contribution is small compared to the other parameters, so these were not included in the interpretation of the plot. The loading of AM Rate (M21) and FM Rate (M23) is positive, so the value of these parameters increase along the PC2 axis. All other parameters decrease along the PC2 axis. Moving from top to bottom along the PC2 axis are Squeaks with the lowest values for the temporal parameters, followed by Barks, Trumpets and Roars, then Rumbles, then Squeals with the highest values for temporal parameters.

Trumpets and Barks show the least within-call variation along the PC2 axis, so these call types are most stereotypic in their temporal pattern. Squeaks and Barks are the shortest calls. Squeals and Rumbles include the longest calls and show the greatest degree of variation along the PC2 axis, so these call types are highly variable in their temporal pattern. These results are consistent with perceptual aural cues of at least duration.

Five of the 6 parameters used as decision rule in the classification were contributors to the principal components. Only Up Sweep Fraction (M28) was not a contributor, and this parameter was used only to differentiate a small number of Squeak calls. In the classification tree, Lower Frequency (M3), Upper Frequency (M4), and FM Rate Variation (M24) separated most of the higher frequency calls (Squeak, Squeal, Trumpet) from the lower frequency calls (Bark, Roar, Rumble), which is consistent with the separation along the PC1 axis. In the classification tree, Duration (M5) separated Barks from the other low frequency calls, Roars and

Rumbles. Roars and Rumbles were separated by Frequency of Peak Overall Intensity (M20), which is supported by the separation along the PC1 axis.

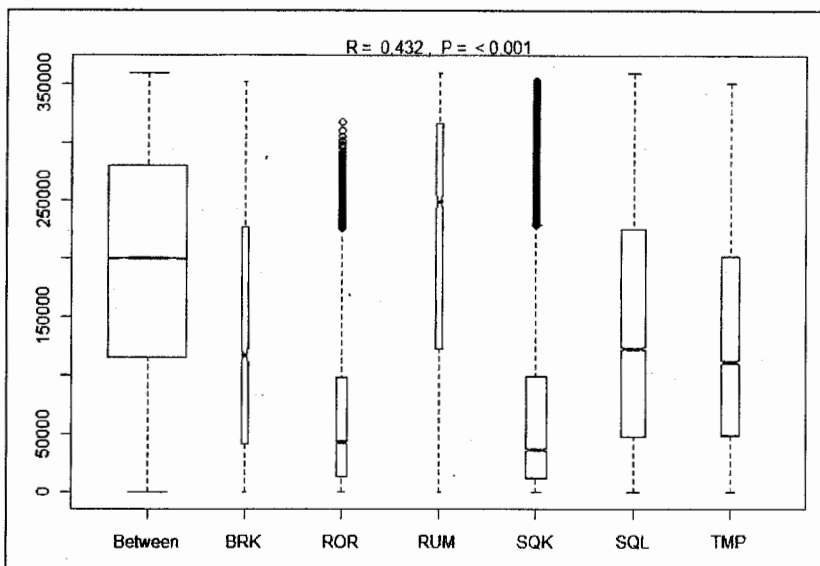
### **Analysis of Similarity (ANOSIM)**

The classification showed a high rate of successful classification, and the parameters that separated call types were consistent with the parameters that contributed to the variance in the Principal Component Analysis, but neither of these analyses determine if the difference among call types is significant. The null hypothesis states that there are no differences in acoustic parameters among the six basic call types (Bark, Roar, Rumble, Squeak, Squeal, Trumpet). A test statistic of differences among groups was computed using an Analysis of Similarity (ANOSIM). Global and pair-wise ANOSIM tests of the scaled dataset were run to determine if there were significant differences in the acoustic parameters among the pre-defined call types and between pairs of call types. Squeak and Squeal and Roar and Rumble were sometimes challenging to differentiate using perceptual aural cues, and the grouping in the PCA showed overlap of these call types. ANOSIM tests were used to determine if these call types were significantly different or if they were gradations of a composite call type.

The global ANOSIM test showed significant differences among the call types ( $R=0.432$ ,  $p<0.001$ ), as shown in Figure 13. The pair-wise ANOSIM tests with a Bonferroni-corrected alpha ( $\alpha=0.0033$ ) indicated a significant difference between every call type, as shown in Table 9. Not surprisingly, the dissimilarity was greatest between the Rumble and the higher frequency calls, Squeak ( $R=0.807$ ,  $p<0.001$ ),

Trumpet ( $R=0.731$ ,  $p<0.001$ ) and Squeal ( $R=0.675$ ,  $p<0.001$ ). The call types that are challenging to distinguish by aural cues had mid-range R values, namely Squeal and Squeak ( $R=0.425$ ,  $p<0.001$ ) and Roar and Rumble ( $R=0.417$ ,  $p<0.001$ ), so it appears that these call types are indeed distinct. The most similar pairs were Bark and Roar ( $R=0.227$ ,  $p<0.001$ ), Bark and Rumble ( $R=0.278$ ,  $p<0.001$ ), and Squeal and Trumpet ( $R=0.230$ ,  $p<0.001$ ), which is consistent with misclassification rates in the classification trees and the clustering overlap in the PCA plots. These calls were easy to distinguish by aural cues, so it may be that the acoustic structure is more similar across acoustic features that are difficult to perceive.

It is important to note that the group sizes were not balanced, so the statistical power is low, and the pairs with low R values may not be significantly different in a dataset with balanced group sizes.



**Figure 13: Global ANOSIM test for difference in parameters among the call types.** TMP=Trumpet, RUM=Rumble, SQK=Squeak, SQL=Squeal, BLW=Blow, BRK=Bark, ROR=Roar Plot shows that the difference between call types is larger than the difference within each call type.

**Table 9: Summary of pair-wise ANOSIM showing significant difference between all call types**

Call pair	R value	Significance	Significant Difference (Bonferroni-corrected $\alpha=0.0033$ )
BRK v. ROR	0.227	<0.001	Yes
BRK v. RUM	0.278	<0.001	Yes
BRK v. SQK	0.539	<0.001	Yes
BRK v. SQL	0.448	<0.001	Yes
BRK v. TMP	0.342	=0.001	Yes
ROR v. RUM	0.417	<0.001	Yes
ROR v. SQK	0.685	<0.001	Yes
ROR v. SQL	0.453	<0.001	Yes
ROR v. TMP	0.468	<0.001	Yes
RUM v. SQK	0.807	<0.001	Yes
RUM v. SQL	0.675	<0.001	Yes
RUM v. TMP	0.731	<0.001	Yes
SQK v. SQL	0.390	<0.001	Yes
SQK v. TMP	0.422	<0.001	Yes
SQL v. TMP	0.230	<0.001	Yes

**Descriptions of call types by acoustic parameters**

Acoustic measurements using Osprey yielded comparative measurements of the basic call types (Bark, Roar, Rumble, Squeak, Squeal, Trumpet) and Blow. The parameters selected for describing the calls types were those that were shown to be most important in separating call types by the classification tree and PCA. Out of the 18 parameters that were included in the analyses, 14 were shown to be important in differentiating call types. Four of the parameters (Bandwidth (M6), Median Frequency (M11), Median Time (M7), and Temporal Concentration (M9)) were highly correlated with other parameters and were not used by the classification tree, but were included in the final description for completeness. Table 10 lists median values for these 14 parameters across the basic call types for quick comparison of acoustic structure.

**Table 10: Median values of acoustic parameters shown by the classification tree and PCA to differentiate basic call types**

Measurements are for the inner 90% of the signal energy rather than the entire signal.

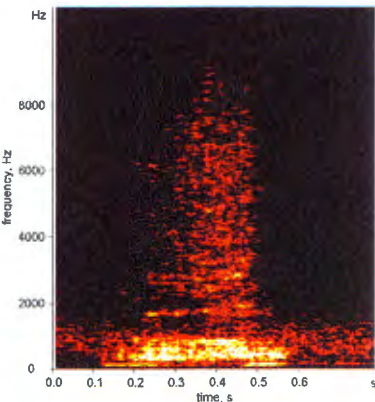
	<b>Bark</b>	<b>Roar</b>	<b>Rumble</b>	<b>Squeak</b>	<b>Squeal</b>	<b>Trumpet</b>
Lower Frequency (Hz) (M3)	75	124	26	495	1031	484
Upper Frequency (Hz) (M4)	1065	777	665	1830	2896	4620
Duration (sec) (M5)	0.5	1.5	2.3	0.2	1.9	0.5
Bandwidth (Hz) (M6)	991	657	583	1318	1873	4091
Median Time (sec) (M7)	0.22	0.66	0.98	0.10	0.75	0.24
Temporal Concentration (sec) (M9)	0.22	0.74	1.21	0.10	1.07	0.27
Median Frequency (Hz) (M11)	408	322	165	912	1509	1305
Frequency of Peak Overall Intensity (Hz) (M20)	172	237	133	775	1443	1098
AM Rate (Hz) (M21)	2.0	0.9	0.5	4.7	2.3	2.2
AM Rate Variation (Hz) (M22)	0.029	0.046	0.372	0.012	0.023	0.023
FM Rate (Hz) (M23)	2.3	0.8	0.5	4.9	1.5	2.7
FM Rate Variation (Hz) (M24)	0.006	0.012	0.093	0.003	0.006	0.006
Overall Entropy (Hz) (M26)	91	42	25	136	145	184
Upsweep Fraction (%) (M28)	46	50	44	34	47	49

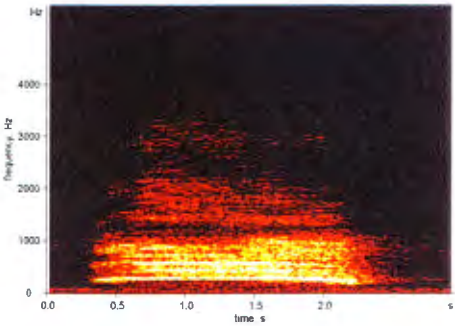
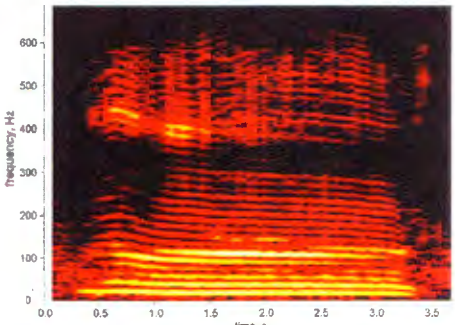
Table 11 provides the final acoustic ethogram for the basic call types (Bark, Roar, Rumble, Squeak, Squeal, Trumpet) plus Blow. This ethogram describes calls both by aural cues and acoustic parameters. Values for Median Time (M7) and Temporal Concentration (M9) in relation to Duration (M5) showed that energy was

distributed fairly evenly around the center of the call for all call types, so these values were not included in the final ethogram. What we can say instead is that the energy is distributed around the center of the call for all of the basic call types in the repertoire. Upsweep Fraction (M28) was not included in the description because it differentiates only a small number of Squeaks.

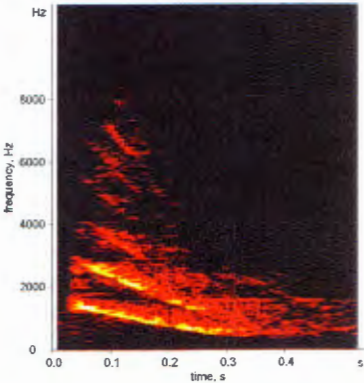
All Osprey measurements are made on the upper 90% of the call energy, so range values may be greater for calls measured with signal processing tools that measure extreme values.

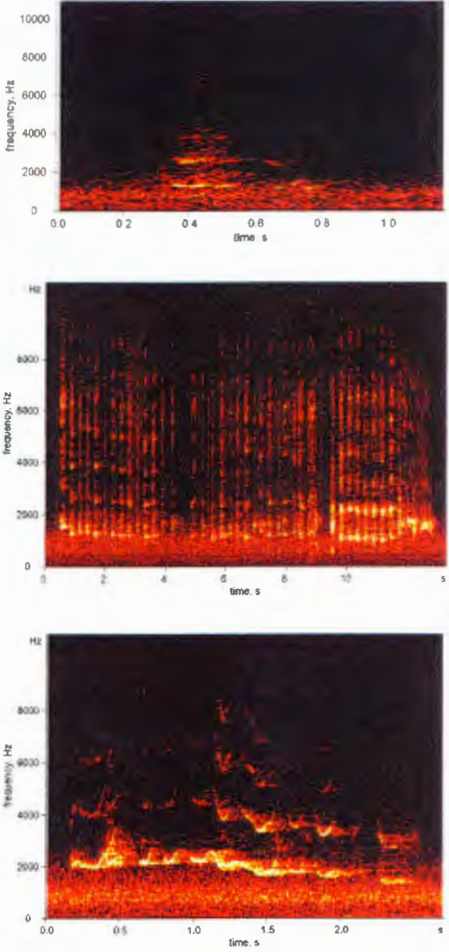
**Table 11: Final ethogram for Bark, Roar, Rumble, Squeak, Squeal, Trumpet, and Blow**  
 Measurements are for the inner 90% of the signal energy rather than the entire signal. The spectrograms are not the same frequency range or duration. The images were captures to show the range of the particular call type.

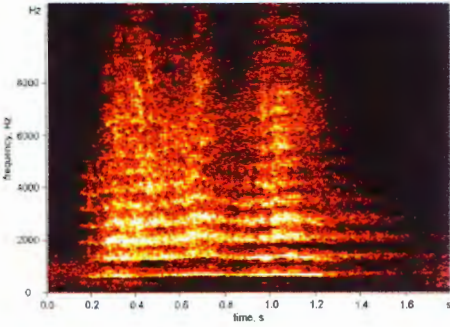
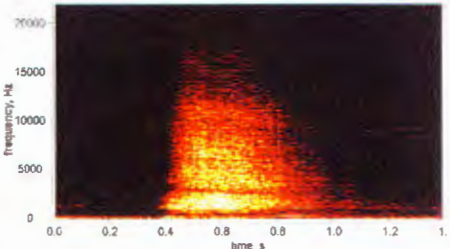
<i>Call type and spectrogram</i>	<i>Definition</i>																						
<p style="text-align: center;"><b>BARK</b></p> 	<p><i>Aural cue:</i> Sounds like a short loud grunt, clearing throat, cape buffalo, chuff of cow, cheetah. Ends abruptly with no roll off. Noisy compared to rumble.</p> <p><i>Visual cue of caller:</i> Mouth open.</p> <p><i>Measurements of 90% signal energy</i></p> <table border="0"> <tr> <td>Duration</td> <td>0.1 – 1.3 sec (median 0.5)</td> </tr> <tr> <td>Lower Frequency</td> <td>26 – 240 Hz (median 75)</td> </tr> <tr> <td>Upper Frequency</td> <td>248 – 4963 Hz (median 1065)</td> </tr> <tr> <td>Bandwidth</td> <td>129 – 4910 Hz (median 991)</td> </tr> <tr> <td>Median Frequency</td> <td>101– 1182 Hz (median 408)</td> </tr> <tr> <td>Freq of Peak Intensity</td> <td>37 – 840 Hz (median 172)</td> </tr> <tr> <td>AM Rate</td> <td>0 – 21.5 Hz (median 2.0)</td> </tr> <tr> <td>AM Rate Variation</td> <td>0 – 0.186 Hz (median 0.029)</td> </tr> <tr> <td>FM Rate</td> <td>0 – 24.8 Hz (median 2.3)</td> </tr> <tr> <td>FM Rate Variation</td> <td>0 – 0.046 Hz (median 0.006)</td> </tr> <tr> <td>Overall Entropy</td> <td>17 – 468 (median 91)</td> </tr> </table>	Duration	0.1 – 1.3 sec (median 0.5)	Lower Frequency	26 – 240 Hz (median 75)	Upper Frequency	248 – 4963 Hz (median 1065)	Bandwidth	129 – 4910 Hz (median 991)	Median Frequency	101– 1182 Hz (median 408)	Freq of Peak Intensity	37 – 840 Hz (median 172)	AM Rate	0 – 21.5 Hz (median 2.0)	AM Rate Variation	0 – 0.186 Hz (median 0.029)	FM Rate	0 – 24.8 Hz (median 2.3)	FM Rate Variation	0 – 0.046 Hz (median 0.006)	Overall Entropy	17 – 468 (median 91)
Duration	0.1 – 1.3 sec (median 0.5)																						
Lower Frequency	26 – 240 Hz (median 75)																						
Upper Frequency	248 – 4963 Hz (median 1065)																						
Bandwidth	129 – 4910 Hz (median 991)																						
Median Frequency	101– 1182 Hz (median 408)																						
Freq of Peak Intensity	37 – 840 Hz (median 172)																						
AM Rate	0 – 21.5 Hz (median 2.0)																						
AM Rate Variation	0 – 0.186 Hz (median 0.029)																						
FM Rate	0 – 24.8 Hz (median 2.3)																						
FM Rate Variation	0 – 0.046 Hz (median 0.006)																						
Overall Entropy	17 – 468 (median 91)																						

Call type and spectrogram	Definition																						
<p style="text-align: center;"><b>ROAR</b></p> 	<p><i>Aural cue:</i> Sounds like a lion roar with changing pitch. Pitch sounds like it changes because of amplitude modulation. Perceived as low frequency.</p> <p><i>Visual cue of caller:</i> Mouth wide open or widens during call, lower abdomen contracts.</p> <p><i>Measurements of 90% signal energy</i></p> <table border="0"> <tr> <td>Duration</td> <td>0.5 – 3.7 sec (median 1.5)</td> </tr> <tr> <td>Lower Frequency</td> <td>23 – 361 Hz (median 124)</td> </tr> <tr> <td>Upper Frequency</td> <td>210 – 3763 Hz (median 777)</td> </tr> <tr> <td>Bandwidth</td> <td>54 – 3650 Hz (median 657)</td> </tr> <tr> <td>Median Frequency</td> <td>122 – 1214 Hz (median 322)</td> </tr> <tr> <td>Freq of Peak Intensity</td> <td>40 – 1292 Hz (median 237)</td> </tr> <tr> <td>AM Rate</td> <td>0.3 – 4.3 Hz (median 0.9)</td> </tr> <tr> <td>AM Rate Variation</td> <td>0.012 – 0.464 Hz (median 0.046)</td> </tr> <tr> <td>FM Rate</td> <td>0.3 – 6.8 Hz (median 0.8)</td> </tr> <tr> <td>FM Rate Variation</td> <td>0.012 – 0.232 Hz (median 0.012)</td> </tr> <tr> <td>Overall Entropy</td> <td>11 – 208 (median 42)</td> </tr> </table>	Duration	0.5 – 3.7 sec (median 1.5)	Lower Frequency	23 – 361 Hz (median 124)	Upper Frequency	210 – 3763 Hz (median 777)	Bandwidth	54 – 3650 Hz (median 657)	Median Frequency	122 – 1214 Hz (median 322)	Freq of Peak Intensity	40 – 1292 Hz (median 237)	AM Rate	0.3 – 4.3 Hz (median 0.9)	AM Rate Variation	0.012 – 0.464 Hz (median 0.046)	FM Rate	0.3 – 6.8 Hz (median 0.8)	FM Rate Variation	0.012 – 0.232 Hz (median 0.012)	Overall Entropy	11 – 208 (median 42)
Duration	0.5 – 3.7 sec (median 1.5)																						
Lower Frequency	23 – 361 Hz (median 124)																						
Upper Frequency	210 – 3763 Hz (median 777)																						
Bandwidth	54 – 3650 Hz (median 657)																						
Median Frequency	122 – 1214 Hz (median 322)																						
Freq of Peak Intensity	40 – 1292 Hz (median 237)																						
AM Rate	0.3 – 4.3 Hz (median 0.9)																						
AM Rate Variation	0.012 – 0.464 Hz (median 0.046)																						
FM Rate	0.3 – 6.8 Hz (median 0.8)																						
FM Rate Variation	0.012 – 0.232 Hz (median 0.012)																						
Overall Entropy	11 – 208 (median 42)																						
<p style="text-align: center;"><b>RUMBLE</b></p> 	<p><i>Aural cue:</i> Sounds like a diesel engine, motorboat, distant helicopter. Pulsating with a flat pitch. Perceived as low frequency.</p> <p><i>Visual cue of caller:</i> Mouth visibly open, forehead flutter.</p> <p><i>Measurements of 90% signal energy</i></p> <table border="0"> <tr> <td>Duration</td> <td>0.9 – 10.0 sec (median 2.3)</td> </tr> <tr> <td>Lower Frequency</td> <td>14 – 120 Hz (median 26)</td> </tr> <tr> <td>Upper Frequency</td> <td>104 – 2498 Hz (median 665)</td> </tr> <tr> <td>Bandwidth</td> <td>86 – 2422 Hz (median 583)</td> </tr> <tr> <td>Median Frequency</td> <td>26 – 590 Hz (median 165)</td> </tr> <tr> <td>Freq of Peak Intensity</td> <td>17 – 750 Hz (median 133)</td> </tr> <tr> <td>AM Rate</td> <td>0.1 – 3.6 Hz (median 0.5)</td> </tr> <tr> <td>AM Rate Variation</td> <td>0.023 – 2.972 Hz (median 0.372)</td> </tr> <tr> <td>FM Rate</td> <td>0.1 – 2.7 Hz (median 0.5)</td> </tr> <tr> <td>FM Rate Variation</td> <td>0.023 – 0.929 Hz (median 0.093)</td> </tr> <tr> <td>Overall Entropy</td> <td>4 – 133 Hz (median 25)</td> </tr> </table>	Duration	0.9 – 10.0 sec (median 2.3)	Lower Frequency	14 – 120 Hz (median 26)	Upper Frequency	104 – 2498 Hz (median 665)	Bandwidth	86 – 2422 Hz (median 583)	Median Frequency	26 – 590 Hz (median 165)	Freq of Peak Intensity	17 – 750 Hz (median 133)	AM Rate	0.1 – 3.6 Hz (median 0.5)	AM Rate Variation	0.023 – 2.972 Hz (median 0.372)	FM Rate	0.1 – 2.7 Hz (median 0.5)	FM Rate Variation	0.023 – 0.929 Hz (median 0.093)	Overall Entropy	4 – 133 Hz (median 25)
Duration	0.9 – 10.0 sec (median 2.3)																						
Lower Frequency	14 – 120 Hz (median 26)																						
Upper Frequency	104 – 2498 Hz (median 665)																						
Bandwidth	86 – 2422 Hz (median 583)																						
Median Frequency	26 – 590 Hz (median 165)																						
Freq of Peak Intensity	17 – 750 Hz (median 133)																						
AM Rate	0.1 – 3.6 Hz (median 0.5)																						
AM Rate Variation	0.023 – 2.972 Hz (median 0.372)																						
FM Rate	0.1 – 2.7 Hz (median 0.5)																						
FM Rate Variation	0.023 – 0.929 Hz (median 0.093)																						
Overall Entropy	4 – 133 Hz (median 25)																						



Call type and spectrogram	Definition																						
<p style="text-align: center;"><b>SQUEAK</b></p> 	<p><i>Aural cue:</i> Sounds like a wet finger across tight rubber surface. audible pedestrian crossing signal, baby alligator, gibbon call. High-pitched sound that falls in pitch, powerful initial sound.</p> <p><i>Visual cue of caller:</i> Cheeks depressed. mouth open slightly or closed, tail sometimes erect.</p> <p><i>Measurements of 90% signal energy</i></p> <table border="0"> <tr> <td>Duration</td> <td>0.1 – 1.2 sec (median 0.2)</td> </tr> <tr> <td>Lower Frequency</td> <td>194 – 883 Hz (median 495)</td> </tr> <tr> <td>Upper Frequency</td> <td>797 – 3768 Hz (median 1830)</td> </tr> <tr> <td>Bandwidth</td> <td>301 – 3273 Hz (median 1318)</td> </tr> <tr> <td>Median Frequency</td> <td>538 – 1905 Hz (median 912)</td> </tr> <tr> <td>Freq of Peak Intensity</td> <td>537– 2024 Hz (median 775)</td> </tr> <tr> <td>AM Rate</td> <td>0.8 – 44.9 Hz (median 4.7)</td> </tr> <tr> <td>AM Rate Variation</td> <td>0.003 – 0.026 Hz (median 0.0012)</td> </tr> <tr> <td>FM Rate</td> <td>1.1 – 70.5 Hz (median 4.9)</td> </tr> <tr> <td>FM Rate Variation</td> <td>0.003 – 0.032 Hz (median 0.003)</td> </tr> <tr> <td>Overall Entropy</td> <td>58 – 294 Hz (median 136)</td> </tr> </table>	Duration	0.1 – 1.2 sec (median 0.2)	Lower Frequency	194 – 883 Hz (median 495)	Upper Frequency	797 – 3768 Hz (median 1830)	Bandwidth	301 – 3273 Hz (median 1318)	Median Frequency	538 – 1905 Hz (median 912)	Freq of Peak Intensity	537– 2024 Hz (median 775)	AM Rate	0.8 – 44.9 Hz (median 4.7)	AM Rate Variation	0.003 – 0.026 Hz (median 0.0012)	FM Rate	1.1 – 70.5 Hz (median 4.9)	FM Rate Variation	0.003 – 0.032 Hz (median 0.003)	Overall Entropy	58 – 294 Hz (median 136)
Duration	0.1 – 1.2 sec (median 0.2)																						
Lower Frequency	194 – 883 Hz (median 495)																						
Upper Frequency	797 – 3768 Hz (median 1830)																						
Bandwidth	301 – 3273 Hz (median 1318)																						
Median Frequency	538 – 1905 Hz (median 912)																						
Freq of Peak Intensity	537– 2024 Hz (median 775)																						
AM Rate	0.8 – 44.9 Hz (median 4.7)																						
AM Rate Variation	0.003 – 0.026 Hz (median 0.0012)																						
FM Rate	1.1 – 70.5 Hz (median 4.9)																						
FM Rate Variation	0.003 – 0.032 Hz (median 0.003)																						
Overall Entropy	58 – 294 Hz (median 136)																						

Call type and spectrogram	Definition																						
<p style="text-align: center;"><b>SQUEAL</b></p> 	<p>Highly variable in structure and duration. Variations numbered in aural cues.</p> <p><i>Aural cue:</i> (1) Short utterance. Sounds like rubber shoes on a gym floor, release of air out of a balloon, windshield wiper on a dry window, a dog whine. (2) Train of repeated Squeal with short breaks between sounds. Sounds like squeals in staccato, intermittent release of air from a balloon. Each squeal sounds flat in pitch but they jump up and down. (3) Sounds like continual release of air from a balloon, fingers on a wet balloon, sliding pitch, sounds long.</p> <p><i>Visual cue of caller:</i> Cheeks depressed, mouth open slightly or closed.</p> <p><i>Measurements of 90% signal energy</i></p> <table border="0"> <tr> <td>Duration</td> <td>0.1 – 16.5 sec (median 1.9)</td> </tr> <tr> <td>Lower Frequency</td> <td>226 – 1927 Hz (median 1031)</td> </tr> <tr> <td>Upper Frequency</td> <td>851 – 8193 Hz (median 2896)</td> </tr> <tr> <td>Bandwidth</td> <td>129 – 7816 Hz (median 1873)</td> </tr> <tr> <td>Median Frequency</td> <td>575 – 2384 Hz (median 1509)</td> </tr> <tr> <td>Freq of Peak Intensity</td> <td>422 – 2218 Hz (median 1443)</td> </tr> <tr> <td>AM Rate</td> <td>0.1 – 12.3 Hz (median 2.3)</td> </tr> <tr> <td>AM Rate Variation</td> <td>0.006 – 0.087 Hz (median 0.023)</td> </tr> <tr> <td>FM Rate</td> <td>0.1 – 23.6 (median 1.5)</td> </tr> <tr> <td>FM Rate Variation</td> <td>0.006– 0.041 (median 0.006)</td> </tr> <tr> <td>Overall Entropy</td> <td>35 – 479 (median 145)</td> </tr> </table>	Duration	0.1 – 16.5 sec (median 1.9)	Lower Frequency	226 – 1927 Hz (median 1031)	Upper Frequency	851 – 8193 Hz (median 2896)	Bandwidth	129 – 7816 Hz (median 1873)	Median Frequency	575 – 2384 Hz (median 1509)	Freq of Peak Intensity	422 – 2218 Hz (median 1443)	AM Rate	0.1 – 12.3 Hz (median 2.3)	AM Rate Variation	0.006 – 0.087 Hz (median 0.023)	FM Rate	0.1 – 23.6 (median 1.5)	FM Rate Variation	0.006– 0.041 (median 0.006)	Overall Entropy	35 – 479 (median 145)
Duration	0.1 – 16.5 sec (median 1.9)																						
Lower Frequency	226 – 1927 Hz (median 1031)																						
Upper Frequency	851 – 8193 Hz (median 2896)																						
Bandwidth	129 – 7816 Hz (median 1873)																						
Median Frequency	575 – 2384 Hz (median 1509)																						
Freq of Peak Intensity	422 – 2218 Hz (median 1443)																						
AM Rate	0.1 – 12.3 Hz (median 2.3)																						
AM Rate Variation	0.006 – 0.087 Hz (median 0.023)																						
FM Rate	0.1 – 23.6 (median 1.5)																						
FM Rate Variation	0.006– 0.041 (median 0.006)																						
Overall Entropy	35 – 479 (median 145)																						

Call type and spectrogram	Definition																						
<p style="text-align: center;"><b>TRUMPET</b></p> 	<p><i>Aural cue:</i> Sounds like a trumpet blast. Perceived as a higher frequency call.</p> <p><i>Visual cue of caller:</i> maybe a lifting at base of trunk, trunk sometimes extended, tail sometimes erect sometimes (any behavior indicating extreme arousal).</p> <p><i>Measurements of 90% signal energy</i></p> <table border="0"> <tr> <td>Duration</td> <td>0.1 – 2.4 sec (median 0.5)</td> </tr> <tr> <td>Lower Frequency</td> <td>54 – 980 Hz (median 485)</td> </tr> <tr> <td>Upper Frequency</td> <td>1281 – 8926 Hz (median 4620)</td> </tr> <tr> <td>Bandwidth</td> <td>1012 – 8742 Hz (median 4091)</td> </tr> <tr> <td>Median Frequency</td> <td>146 – 4454 Hz (median 1305)</td> </tr> <tr> <td>Freq of Peak Intensity</td> <td>129 – 3531 Hz (median 1098)</td> </tr> <tr> <td>AM Rate</td> <td>0.4 – 27.6 Hz (median 2.2)</td> </tr> <tr> <td>AM Rate Variation</td> <td>0.006 – 0.061 Hz (median 0.023)</td> </tr> <tr> <td>FM Rate</td> <td>0.4 – 23.1 Hz (median 2.7)</td> </tr> <tr> <td>FM Rate Variation</td> <td>0.006 – 0.244 Hz (median 0.006)</td> </tr> <tr> <td>Overall Entropy</td> <td>29 – 831 Hz (median 184)</td> </tr> </table>	Duration	0.1 – 2.4 sec (median 0.5)	Lower Frequency	54 – 980 Hz (median 485)	Upper Frequency	1281 – 8926 Hz (median 4620)	Bandwidth	1012 – 8742 Hz (median 4091)	Median Frequency	146 – 4454 Hz (median 1305)	Freq of Peak Intensity	129 – 3531 Hz (median 1098)	AM Rate	0.4 – 27.6 Hz (median 2.2)	AM Rate Variation	0.006 – 0.061 Hz (median 0.023)	FM Rate	0.4 – 23.1 Hz (median 2.7)	FM Rate Variation	0.006 – 0.244 Hz (median 0.006)	Overall Entropy	29 – 831 Hz (median 184)
Duration	0.1 – 2.4 sec (median 0.5)																						
Lower Frequency	54 – 980 Hz (median 485)																						
Upper Frequency	1281 – 8926 Hz (median 4620)																						
Bandwidth	1012 – 8742 Hz (median 4091)																						
Median Frequency	146 – 4454 Hz (median 1305)																						
Freq of Peak Intensity	129 – 3531 Hz (median 1098)																						
AM Rate	0.4 – 27.6 Hz (median 2.2)																						
AM Rate Variation	0.006 – 0.061 Hz (median 0.023)																						
FM Rate	0.4 – 23.1 Hz (median 2.7)																						
FM Rate Variation	0.006 – 0.244 Hz (median 0.006)																						
Overall Entropy	29 – 831 Hz (median 184)																						
<p style="text-align: center;"><b>BLOW</b></p> 	<p><i>Aural cue:</i> Forced exhalation through end of trunk. Not a vocalization.</p> <p><i>Visual cue of caller:</i> air blast from trunk</p> <p><i>Measurements of 90% signal energy</i></p> <table border="0"> <tr> <td>Duration</td> <td>0.3 – 1.2 sec (median 0.7)</td> </tr> <tr> <td>Lower Frequency</td> <td>81 – 1512 Hz (median 372)</td> </tr> <tr> <td>Upper Frequency</td> <td>1448 – 9318 Hz (median 3790)</td> </tr> <tr> <td>Bandwidth</td> <td>1217 – 8990 Hz (median 3273)</td> </tr> <tr> <td>Median Frequency</td> <td>531 – 4021 Hz (median 1156)</td> </tr> <tr> <td>Freq of Peak Intensity</td> <td>97 – 3295 Hz (median 894)</td> </tr> <tr> <td>AM Rate</td> <td>0.8 – 3.6 Hz (median 1.4)</td> </tr> <tr> <td>AM Rate Variation</td> <td>0.012 – 0.07 Hz (median 0.046)</td> </tr> <tr> <td>FM Rate</td> <td>0.8 – 25.1 Hz (median 1.6)</td> </tr> <tr> <td>FM Rate Variation</td> <td>0.012 – 0.046 Hz (median 0.012)</td> </tr> <tr> <td>Overall Entropy</td> <td>102 – 875 Hz (median 286)</td> </tr> </table>	Duration	0.3 – 1.2 sec (median 0.7)	Lower Frequency	81 – 1512 Hz (median 372)	Upper Frequency	1448 – 9318 Hz (median 3790)	Bandwidth	1217 – 8990 Hz (median 3273)	Median Frequency	531 – 4021 Hz (median 1156)	Freq of Peak Intensity	97 – 3295 Hz (median 894)	AM Rate	0.8 – 3.6 Hz (median 1.4)	AM Rate Variation	0.012 – 0.07 Hz (median 0.046)	FM Rate	0.8 – 25.1 Hz (median 1.6)	FM Rate Variation	0.012 – 0.046 Hz (median 0.012)	Overall Entropy	102 – 875 Hz (median 286)
Duration	0.3 – 1.2 sec (median 0.7)																						
Lower Frequency	81 – 1512 Hz (median 372)																						
Upper Frequency	1448 – 9318 Hz (median 3790)																						
Bandwidth	1217 – 8990 Hz (median 3273)																						
Median Frequency	531 – 4021 Hz (median 1156)																						
Freq of Peak Intensity	97 – 3295 Hz (median 894)																						
AM Rate	0.8 – 3.6 Hz (median 1.4)																						
AM Rate Variation	0.012 – 0.07 Hz (median 0.046)																						
FM Rate	0.8 – 25.1 Hz (median 1.6)																						
FM Rate Variation	0.012 – 0.046 Hz (median 0.012)																						
Overall Entropy	102 – 875 Hz (median 286)																						

## **Nonlinear dynamics of vocal production**

These basic call types exhibited some nonlinear dynamics of vocal production, which include subharmonics, deterministic chaos, bifurcations, biphonation, and frequency jumps (Mann et al., 2006). Subharmonics are sounds at frequencies other than the fundamental frequency or integer harmonic of the fundamental that can be generated by coupled oscillators. Deterministic chaos comprises noisy signals that are not random noise, but generated by chaotic process. Bifurcation is the rapid shifts between tonal harmonics and chaos. Biphonation is the generation of two independent frequencies simultaneously. Frequency jumps are abrupt up or down transitions between two frequencies or harmonics that are unpredictable (Mann et al., 2006). The Trumpet exhibited clear evidence of bifurcation, as measured by the Overall Entropy (M26) parameter and by visual inspection of Trumpet spectrograms. The Squeak and Squeal also exhibited evidence of bifurcation as measured by the Overall Entropy (M26) parameter, but it was not evident from visual inspection of the spectrogram. The Roar and Bark appear to be noisier sounds, but the shift from linearity to non-linearity is not evident. The Squeal exhibited clear frequency jumps when produced in a train of Squeals. The degree to which these non-linearities vary with perceived level of arousal is being investigated, but is not included here.

## **Comparison of low frequency calls to background noise**

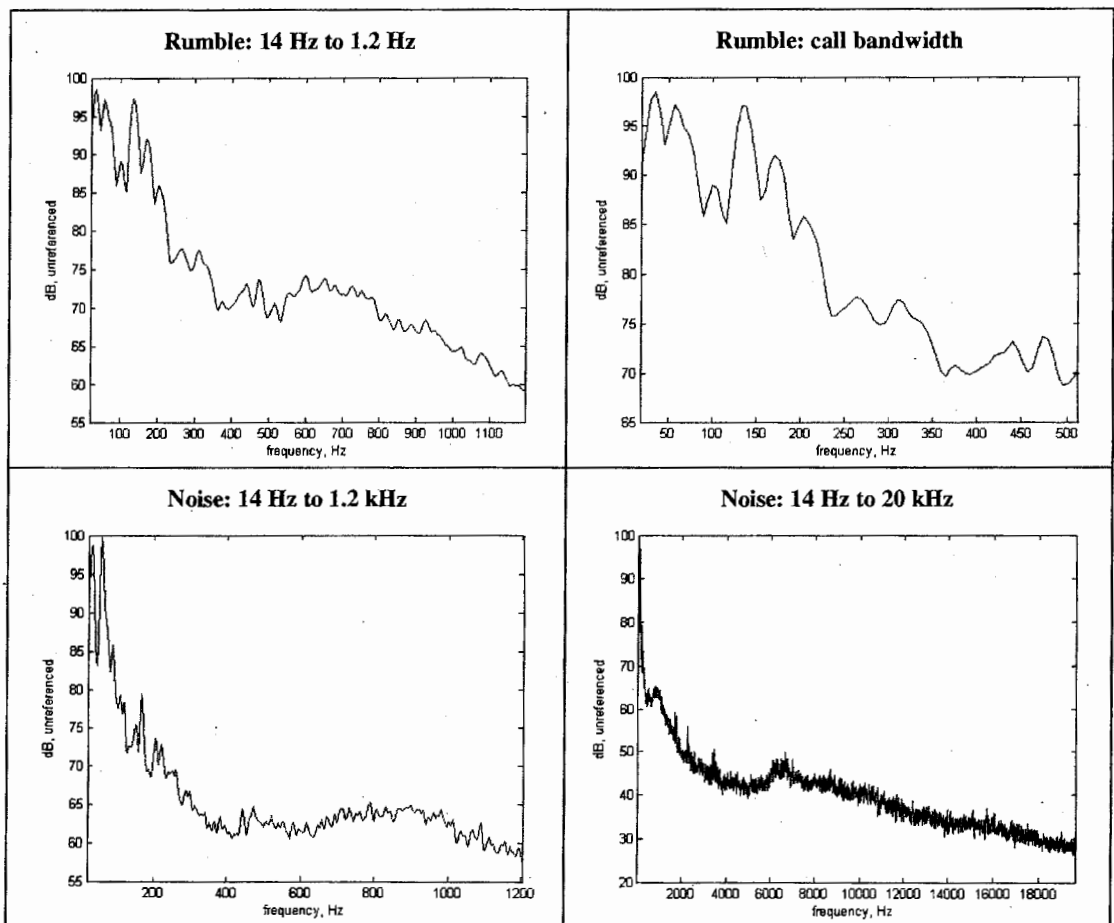
Power spectra of Paired Rumbles and noise for the three study sites were compared to investigate the potential for Rumbles being buried in background noise. At the Oregon Zoo (Figure 14) the Rumble had multiple peaks below 200 Hz. The

noise floor started to rise below 300 Hz, with the peak below 100Hz. From visual inspection, the signal-to-noise ratio was less than 10 dB below 100 Hz, but increased in the 100-300 Hz range. Given this ratio, it appears that signal energy below 100 Hz could become inaudible due to noise in this environment. In the entire range of 15 Hz to 20 kHz, noise declined more sharply from lower frequencies to approximately 2 kHz, then declined gradually or flattened. In one recording there was a slight increase in noise in the 6-8 kHz range.

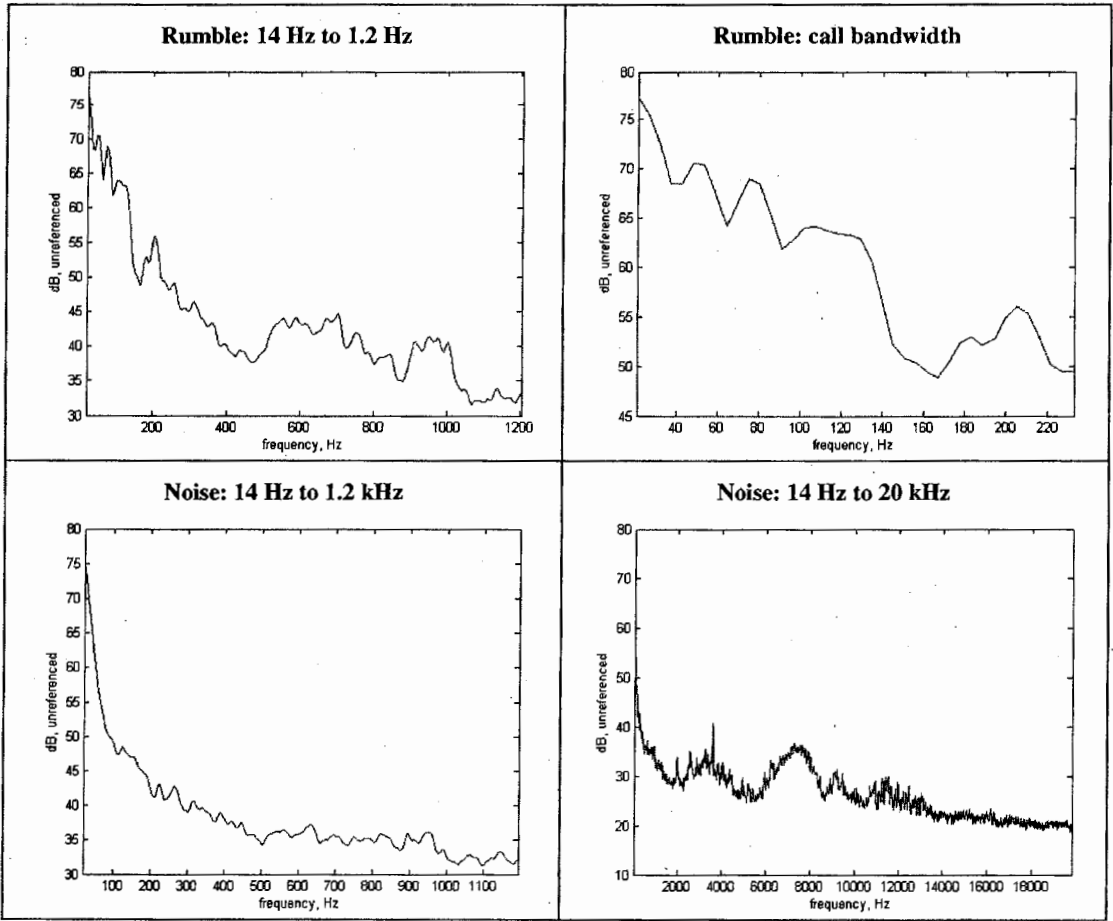
At the Elephant Nature Park in rural Thailand (Figure 15), the Rumble had multiple peaks below 200 Hz. The noise floor started to rise sharply below 100 Hz, with the peak below 100Hz. Noise had the highest power at low frequencies (below approximately 30 Hz). From visual inspection, the signal-to-noise ratio was less than 10 dB below 100 Hz, but increased in the 100-300 Hz range. Given this ratio, it appears that the signal energy below 100 Hz could become inaudible due to noise in this environment as well. Wind gusts were a factor at this site. In the entire range of 15 Hz to 20 kHz, noise declined sharply from low frequency to approximately 400 Hz, then declined gradually across the range. In all selected recordings there was a notable increase in noise in the 2-4 kHz and 6-11 kHz ranges. Based on aural cues at the time of observation, noise in 6-11 kHz band may be sounds of insects or other arthropods.

At the Royal Elephant Kraal in urban Thailand (Figure 16), the Rumble had multiple peaks below 200 Hz. The noise floor started to rise below 300 Hz, with the peak below 100Hz. From visual inspection, the signal-to-noise ratio was less than 10 dB below 200 Hz (except for recording at close proximity), and remained fairly low to

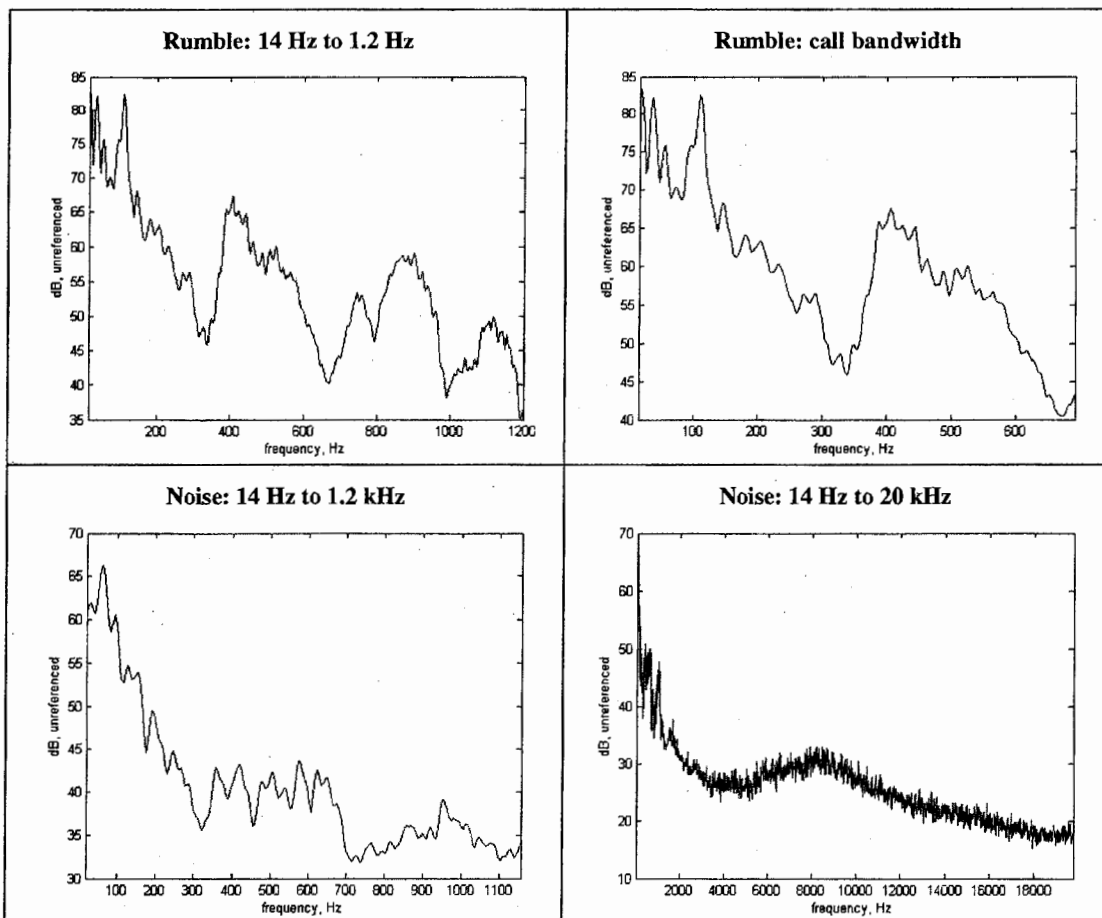
about 300 Hz. Given this ratio, it appears that signal energy below 200 or 300 Hz could become inaudible due to background noise in this environment as well. Wind gusts were a factor at this site. In the entire range of 15 Hz to 20 kHz, noise declined sharply from 15 Hz to approximately 400Hz, then declined gradually across the range. In all selected recordings there was a notable increase in noise in the 6-11 kHz range. With the exception of the increase in the 6-11 kHz range, the noise profile across the entire range of 14 Hz to 20 kHz looked more similar to the Oregon Zoo than to the Elephant Nature Park, so the frequency band of the background noise may be primarily a result of the degree of urbanization.



**Figure 14: Power spectra of Rumble and background noise at the Oregon Zoo**  
 Frequency is on the x-axis. Power (dB) is on the y-axis. Power is referenced to the same value.



**Figure 15: Power spectra of Rumble and background noise at Elephant Nature Park**  
 Frequency is on the x-axis. Power (dB) is on the y-axis. Power is referenced to the same value.



**Figure 16: Power spectra of rumbles and background noise at the Royal Elephant Kraal**  
 Frequency is on the x-axis. Power (dB) is on the y-axis. Power is referenced to the same value.



## CHAPTER V: DISCUSSION

The goals of contributing to the basic science of elephant communication were met by the results of this research. An acoustic repertoire of Asian elephants based on acoustic parameters was defined. A comparison of how the repertoire was used by groups of elephants and individuals was made. The manual methods of classification based on perceptual aural cues and visual inspection of spectrograms was compared to automated classification, and structural distinction among call types was validated. The fact that a limited number of acoustic parameters differentiated call types suggests that these parameters could be used for detection of elephant calls. The next steps would be to determine if the parameter set could be reduced further, and to examine differentiation of elephant sounds from non-elephants, the data of which has already been collected and is ready to analyze. Finally, the descriptions of elephant sounds presented in this thesis serve as a basis for comparisons among captive and wild Asian elephants and between Asian and African elephants.

The final repertoire was defined by 6 basic call types (Bark, Roar, Rumble, Bark, Squeal, Squeal, and Trumpet), 5 call combinations and modifications with these basic types as their constituent parts (Roar-Rumble, Squeal-Squeak, Squeak train, Squeak-Bark, and Trumpet-Roar), and a sound that was produced frequently by many elephants, the Blow. Results suggest these call types are differentiated by 11 temporal and spectral parameters, and with future analyses this feature set may be able to be reduced further.

Given the high success rate of the classification tree using only 6 parameters as decision rules, it appears that this tree does have potential as a predictive model for classifying new calls into pre-defined call types. Consistent with aural cues was the separation of the higher frequency calls (Squeak, Squeal, and Trumpet) from the lower frequency calls (Bark, Roar, Rumble), and the confusion between Roars and Rumbles. Separation of the call types with the Principal Component Analysis appeared to be in agreement with the separation by the classification tree. Finally, analysis of similarity tests showed significant difference among the 6 basic call types and between all pairs of these call types, so these call types are structurally distinct.

Data used in this study were collected in both urban and rural settings with many sources of noise and recorded with multiple recording systems, and results suggest that automated call detection is possible in varying recording situations, with various noise sources and using various recording systems.

Biological sources of variability within call types may include individuality of the caller, social context, level of arousal and state of motivation, and potentially the size of the animals and size of the head space. Some call types appear to be stereotypic in either temporal or spectral structure and other vary widely. The Squeak is a stereotypic call, the sound production of which includes manipulation of the cheek or lips. The Squeal is highly variable, but the sound production appears to be the same. Little is known about sound production in elephants, but it may be that these sounds are produced by some source other than the vibration of vocal cords alone.

Results of repertoire usage agreed with (Langbauer, 2000) in that there was a difference in calling pattern of males and females, with male elephants producing fewer call types than their female counterparts (Langbauer, 2000). The repertoire defined by this study suggests some differences between African elephants and Asian elephants. The most common vocalization of African elephants is the Rumble, which is highly variable and is used in multiple contexts and serves multiple functions (Soltis et al., 2005; Olson, 2004). The Rumble is produced by Asian elephants, but it was not the most common sound produced in this study. The Asian elephants in this study were captive or domesticated and were not wild, but the published calls of wild elephants by McKay (1973) suggest that at least some of the call types defined in this study are also produced by wild Asian elephants. Ecological differences in populations of African and Asian elephants include predator pressures, resource availability, group size, and human encroachment. Home range size may be a function of these differences, and communication modalities and distance may be a function of home range and group size. One could hypothesize that a reason for the Rumble being the primary call of African elephants is a greater need for long distance communication in order to maintain separation while remaining in contact for the purpose of resource utilization.

Future analyses for this dataset include completing the caller identification, investigating the communicative function of these sounds based on behavioral data already collected, investigating potential explanations for differences in call distribution among individuals, running unsupervised call classification as another

validation of established call types, doing a more rigorous analysis of signal-to-noise ratio at various frequency values, and running non-elephant sounds through a classification tree as a preliminary test for call detection potential.

## REFERENCES

- Baken, R. J. 1996. *Clinical Measurement of Speech and Voice*. San Diego: Singular Publishing Group, Inc.
- Blake, S. & Hedges, S. 2004. Sinking the Flagship: the Case of Forest Elephants in Asia and Africa. *Conservation Biology*, **18**, 1191-1202.
- Blanc, J. 2008. *Loxodonta africana*. In: IUCN 2008. 2008 IUCN Red List of Threatened Species. <[www.iucnredlist.org](http://www.iucnredlist.org)>. Downloaded on 11 May 2009.
- Campbell, G. S., Gisiner, R. C., Helweg, D. A. & Milette, L. L. 2002. Acoustic identification of female Steller sea lions (*Eumetopias jubatus*). *The Journal of the Acoustical Society of America*, **111**, 2920-2928.
- Charif, R. A., Ramey, R. R., Langbauer, W. R., Payne, K. B., Martin, R. B. & Brown, L. M. 2005. Spatial relationships and matrilineal kinship in African savanna elephant (*Loxodonta africana*) clans. *Behavioral Ecology and Sociobiology*, **V57**, 327-338.
- Choudhury, A., Lahiri Choudhury, D. K., Desai, A., Duckworth, J. W., Easa, P. S., Johnsingh, A. J. T., Fernando, P., Hedges, S., Gunawardena, M., Kurt, F., Karanth, U., Lister, A., Menon, V., Riddle, H., Rübel, A. & Wikramanayake, E. 2008. *Elephas maximus*. In: IUCN 2008. 2008 IUCN Red List of Threatened Species. <[www.iucnredlist.org](http://www.iucnredlist.org)>. Downloaded on 11 May 2009.
- CITES. 2008. Appendices I, II, and III to Convention on International Trade in Endangered Species of Wild Fauna and Flora <<http://www.cites.org>>. Downloaded on 11 May 2009. Geneva.
- Clemins, P. J., Johnson, M. T., Leong, K. M. & Savage, A. 2005. Automatic classification and speaker identification of African elephant (*Loxodonta africana*) vocalizations. *The Journal of the Acoustical Society of America*, **117**, 956-963.
- Cortopassi, K. A. 2006. Automated and Robust Measurement of Signal Features. Bioacoustics Research Program, Cornell Laboratory of Ornithology, Cornell University.
- Darden, S. K., Dabelsteen, T. & Pedersen, S. B. 2003. A potential tool for swift fox (*Vulpes velox*) conservation: Individuality of long-range barking sequences. *Journal of Mammalogy*, **84**, 1417-1427.
- Deecke, V. B. & Janik, V. M. 2006. Automated categorization of bioacoustic signals: Avoiding perceptual pitfalls. *The Journal of the Acoustical Society of America*, **119**, 645-653.
- Doluweera, G., Fernando, S., Gunasekera, C., Madanayake, A., Rossel, G., Seneviratne, L., Wijetunga, S. & Yapa, S. 2003. Infrasound – Can it be more than a means of communication between Elephants? In: *Symposium on Human-Elephant Relationships and Conflicts*. Sri Lanka.
- Douglas-Hamilton, I. 1972. On the ecology and behaviour of the African elephant. In: *Oriel College*, pp. 257: Oxford.

- Eisenberg, J. F., McKay, G. M. & Jainudeen, M. R. 1971. Reproductive behavior of the Asiatic elephant (*Elephas maximus maximus* L.). *Behaviour*, **38**, 193-225.
- Fernando, P., Vidya, T. N. C., Payne, J., Stuewe, M., Davison, G., Alfred, R. J., Andau, P., Bosi, E., Kilbourn, A. & Melnick, D. J. 2003. DNA analysis indicates that Asian elephants are native to Borneo and are therefore a high priority for conservation. *Plos Biology*, **1**, 110-115.
- Fischer, M. 2004. AZA Elephant TAG/SSP Regional Collection Plan. Saint Louis: Saint Louis Zoo.
- Fitch, W. T. 1997. Vocal tract length and formant frequency dispersion correlate with body size in rhesus macaques. *The Journal of the Acoustical Society of America*, **102**, 1213-1222.
- Fitch, W. T. & Hauser, M. D. 2002. Unpacking "Honesty": Vertebrate Vocal Production and the Evolution of Acoustic Signals. *Springer Handbook of Auditory Research*, **16: Acoustic Communication**, 65-137.
- Fleischer, R. C., Perry, E. A., Muralidharan, K., Stevens, E. E. & Wemmer, C. M. 2001. PHYLOGEOGRAPHY OF THE ASIAN ELEPHANT (*ELEPHAS MAXIMUS*) BASED ON MITOCHONDRIAL DNA. *Evolution*, **55**, 1882-1892.
- Fowler, M. E. & Mikota, S. K. 2006. *Biology, Medicine, and Surgery of Elephants*. Ames: Blackwell Publishing.
- Fristrup, K. M. & Watkins, W. A. 1993. Marine Animal Sound Classification. pp. 29: Woods Hole Oceanographic Institution Technical Report WHOI-94-13.
- Garstang, M. 2004. Long-distance, low-frequency elephant communication. *Journal of Comparative Physiology A: Sensory, Neural, and Behavioral Physiology*, **190**, 791 - 805.
- Garstang, M., Larom, D., Raspel, R. & Lindeque, M. 1995. Atmospheric controls on elephant communication. *The Journal of Experimental Biology*, **198**, 939-951.
- Heffner, R., Heffner, H. & Stichman, N. 1982. Role of the elephant pinna in sound localization. *Animal Behaviour*, **30**, 628-630.
- Heffner, S. R. & Heffner, H. E. 1982. Hearing in the elephant (*Elephas maximus*): absolute sensitivity, frequency discrimination, and sound localization. *Journal of Comparative and Physiological Psychology*, **96**, 926-44.
- IUCN. 2008. 2008 IUCN Red List of Threatened Species. <[www.iucnredlist.org](http://www.iucnredlist.org)>. Downloaded on 11 May 2009.
- Janik, V. M., Sayigh, L. S. & Wells, R. S. 2006. Signature whistle shape conveys identity information to bottlenose dolphins. *Proc Natl Acad Sci U S A*, **103**, 8293-7.
- Kemf, E. & Santiapillai, C. 2000. Asian Elephants in the Wild. In: *WWF Species Status Report* (Ed. by Wilson, A.). Gland, Switzerland: WWF-World Wildlife Fund for Nature.
- Langbauer, J., William R., Payne, K. B., Charif, R. A., Rapaport, L. & Osborn, F. 1991. African elephants respond to distant playbacks of low-frequency conspecific calls. *The Journal of Experimental Biology*, **157**, 35-46.

- Langbauer, J., William R., Payne, K. B., Charif, R. A. & Thomas, E. M. 1989. Responses of captive African elephants to playback of low-frequency calls. *Canadian Journal of Zoology*, **67**, 2604-2607.
- Langbauer, W. R., Jr. 2000. Elephant Communication. *Zoo Biology*, **19**, 425-445.
- Larom, D., Garstang, M., Lindeque, M., Raspet, R., Zunckel, M., Hong, Y., Brassel, K., O'Beirne, S. & Sokolic, F. 1997. Meteorology and elephant infrasound at Etosha National Park, Namibia. *J. Acoust. Soc. Am.*, **101**, 1710-1717.
- Mann, D. A., O' Shea, T. J. & Nowacek, D. P. 2006. Nonlinear Dynamics in Manatee Vocalizations. *Marine Mammal Science*, **22**, 548-555.
- Marter, K., Quine, D. & Marler, P. 1977. Sound transmission and its significance for animal communication II: tropical forest habits. *Behavioral Ecology and Sociobiology*, **2**, 291-302.
- McComb, K. 1996. Studying Vocal Communication in Elephants. In: *Studying Elephants* (Ed. by Kangwana, K.), pp. 112-119. Nairobi, Kenya: African Wildlife Foundation (AWF).
- McComb, K., Baker, D. R. L., Moss, C. & Sayialel, S. 2002. Long-distance communication of social identity in African elephants. *Animal Behaviour*, **65**, 317-329.
- McComb, K., Moss, C., Sayialel, S. & Baker, L. 2000. Unusually extensive networks of vocal recognition in African elephants. *Animal Behaviour*, **59**, 1103-9.
- McComb, K., Reby, D., Baker, L., Moss, C. & Sayialel, S. 2003. Long-distance communication of acoustic cues to social identity in African elephants. *Animal Behaviour*, **66**, 195-195.
- McKay, G. M. 1973. Behavior and ecology of the Asiatic elephant in southeastern Ceylon. In: *Smithsonian Contributions to Zoology*, pp. 1-113. Washington, DC: Smithsonian Institution Press.
- Mellinger, D. K. & Bradbury, J. W. 2007. Acoustic Measurement of Marine Mammal Sounds in Noisy Environments. In: *Second International Conference on Underwater Acoustic Measurements: Technologies and Results*. Heraklion, Greece.
- Moss, C. J. 1983. Oestrous behaviour and female choice in the African elephant. *Behaviour*, **86**, 167-196.
- O'Connell-Rodwell, C., Wood, J., Rodwell, T., Puria, S., Partan, S., Keefe, R., Shriver, D., Arnason, B. & Hart, L. 2006. Wild elephant (*Loxodonta africana*) breeding herds respond to artificially transmitted seismic stimuli. *Behavioral Ecology and Sociobiology*, **59**, 842-850.
- O'Connell-Rodwell, C. E., Wood, J. D., Kinzley, C., Rodwell, T. C., Poole, J. H. & Puria, S. 2007. Wild African elephants (*Loxodonta africana*) discriminate between familiar and unfamiliar conspecific seismic alarm calls. *The Journal of the Acoustical Society of America*, **122**, 823-830.
- Olson, D. 2004. Elephant Husbandry Resource Guide. pp. 280: International Elephant Foundation.

- Payne, K. B., Jr., W. R. L. & Thomas, E. M. 1986. Infrasonic calls of the Asian elephant (*Elephas maximus*). *Behavioral Ecology and Sociobiology*, **18**, 297-301.
- Payne, K. B., Thompson, M. & Kramer, L. 2003. Elephant calling patterns as indicators of group size and composition: The basis for an acoustic monitoring system. *African Journal of Ecology*, **41**, 99-107.
- Poole, J. H. 1987. Rutting behavior in African elephants: the phenomenon of musth. *Behaviour*, **102**, 283-316.
- Poole, J. H. 1989. Announcing intent: the aggressive state of musth in African elephants. *Animal Behaviour*, **37**, 140-152.
- Poole, J. H. 1999. Signals and Assessment in African Elephants: Evidence from playback experiments. *Animal Behaviour*, **58**, 185-193.
- Poole, J. H. in prep. *Elephant vocal communication*. In: *The Amboseli Elephants: A long-term perspective on a long-lived mammal*. Chicago: University of Chicago Press.
- Poole, J. H., Payne, K. B., Jr., W. R. L. & Moss, C. J. 1988. The social contexts of some very low frequency calls of African elephants. *Behavioral Ecology and Sociobiology*, **22**, 385-392.
- Reby, D., Joachim, J., Lauga, L., Lek, L. & Auglagnier, S. 1998. Individuality in the groans of fallow deer (*Dama dama*) bucks. *Journal of Zoology (London)*, **245**, 79-84.
- Rendall, D., Owren, M. J. & Rodman, P. S. 1998. The role of vocal tract filtering in identity cueing in rhesus monkey (*Macaca mulatta*) vocalizations. *The Journal of the Acoustical Society of America*, **103**, 602-614.
- Riddle, H. S., Rasmussen, B. & Schmitt, D. L. 2003. Are captive elephants important to conservation? *Gajah*, **22**, 57-60.
- Roca, A. L., Georgiadis, N., Pecon-Slattery, J. & O'Brien, S. J. 2001. Genetic Evidence for Two Species of Elephant in Africa. *Science*, **293**, 1473-1477.
- Sanvito, S., Galimberti, F. & Miller, E. H. 2007. Vocal signalling of male southern elephant seals is honest but imprecise. *Animal Behaviour*, **73**, 287-299.
- Schulte, B. & Rasmussen, L. 1999. Signal-receiver interplay in the communication of male condition by Asian elephants. *Animal Behaviour*, **57**, 1265-1274.
- SEVP. 2008. Savanna Elephant Vocalization Project  
<<http://www.elephantvoices.org>>.
- Shoshani, J. 2000. *Elephants: majestic creatures of the wild*. New York: Checkmark Books.
- Shoshani, J. & Eisenberg, J. F. 1982. *Elephas maximus*. *Mammalian Species*, 1-8.
- Soltis, J., Leong, K. & Savage, A. 2005. African elephant vocal communication II: rumble variation reflects the individual identity and emotional state of callers. *Animal Behaviour*, **70**, 589-599.
- Stoeger-Horwath, A. S., Stoeger, S., Schwammer, H. M. & Kratochvil, H. 2007. Call repertoire of infant African elephants: First insights into the early vocal ontogeny. *The Journal of the Acoustical Society of America*, **121**, 3922-3931.



- Sukumar, R. 2003. *The living elephants : evolutionary ecology, behavior, and conservation*. New York: Oxford University Press.
- Sukumar, R. 2006. A brief review of the status, distribution and biology of wild Asian elephants *Elephas maximus*. *International Zoo Yearbook*, **40**, 1-8.
- Vidya, T. N. C. & Sukumar, R. 2005. Social organization of the Asian elephant (*Elephas maximus*) in southern India inferred from microsatellite DNA. *Journal of Ethology*, **23**, 205-210.
- von Muggenthaler, E. 1992. Infrasound from the okapi. In: *American Association for the Advancement of Science (A.A.A.S) 158th conference*.
- von Muggenthaler, E. 2000. Infrasonic and low-frequency vocalizations from Siberian and Bengal tigers. *The Journal of the Acoustical Society of America*, **108**, 2541.
- von Muggenthaler, E., Baes, C., Hill, D., Fulk, R. & Lee, A. 2001. Infrasound and low frequency vocalizations from the giraffe; Helmholtz resonance in biology. In: *Acoustical Society of America, regional conference*. South Carolina: Proceedings of Riverbanks Consortium.
- von Muggenthaler, E., Reinhart, P., Lympany, B. & Craft, R. B. 2003. Songlike vocalizations from the Sumatran Rhinoceros (*Dicerorhinus sumatrensis*). *Acoustics Research Letters Online (ARLO)*, *Acoustical Society of America*, **4**, 83-88.

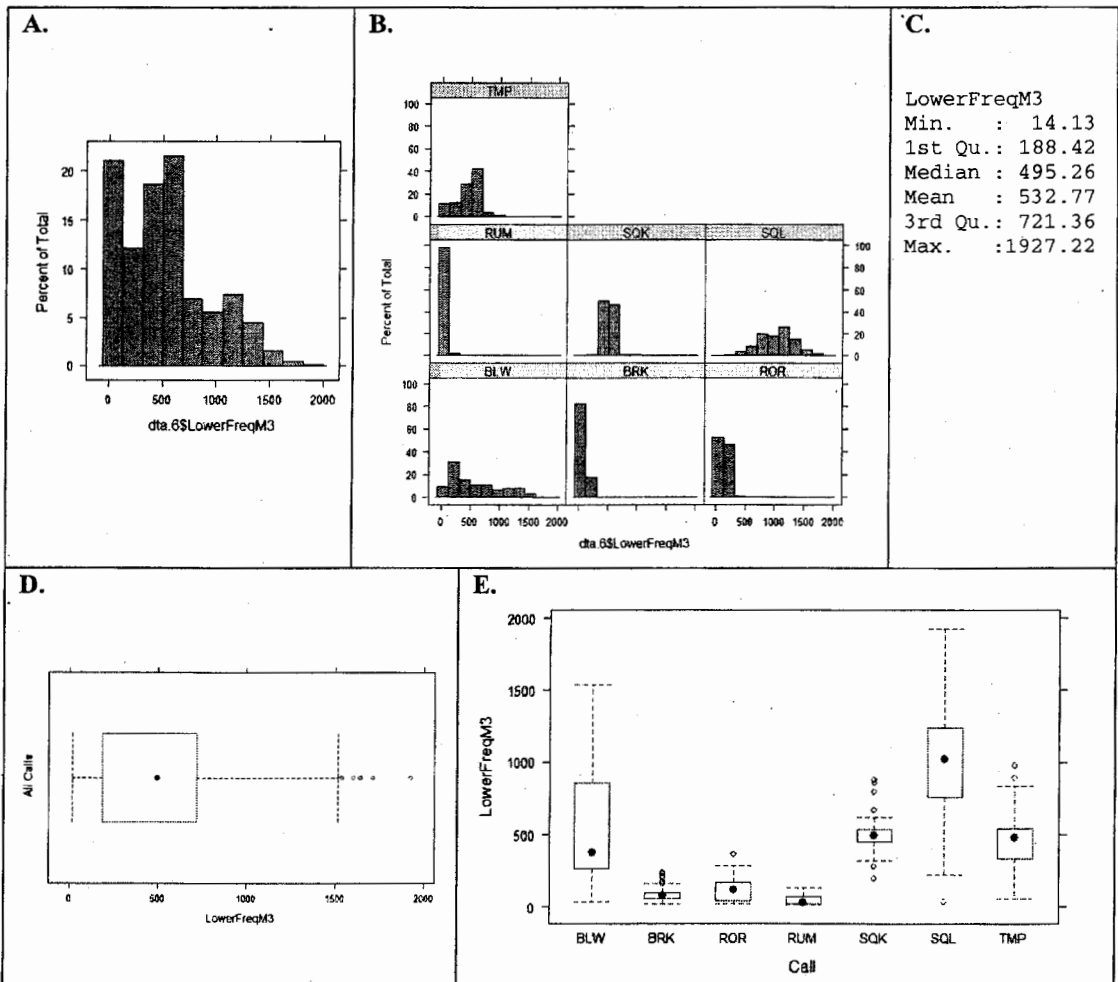
## APPENDIX A: DISTRIBUTION PLOTS OF PARAMETERS

### HISTOGRAMS AND BOXPLOTS OF ORIGINAL DATA

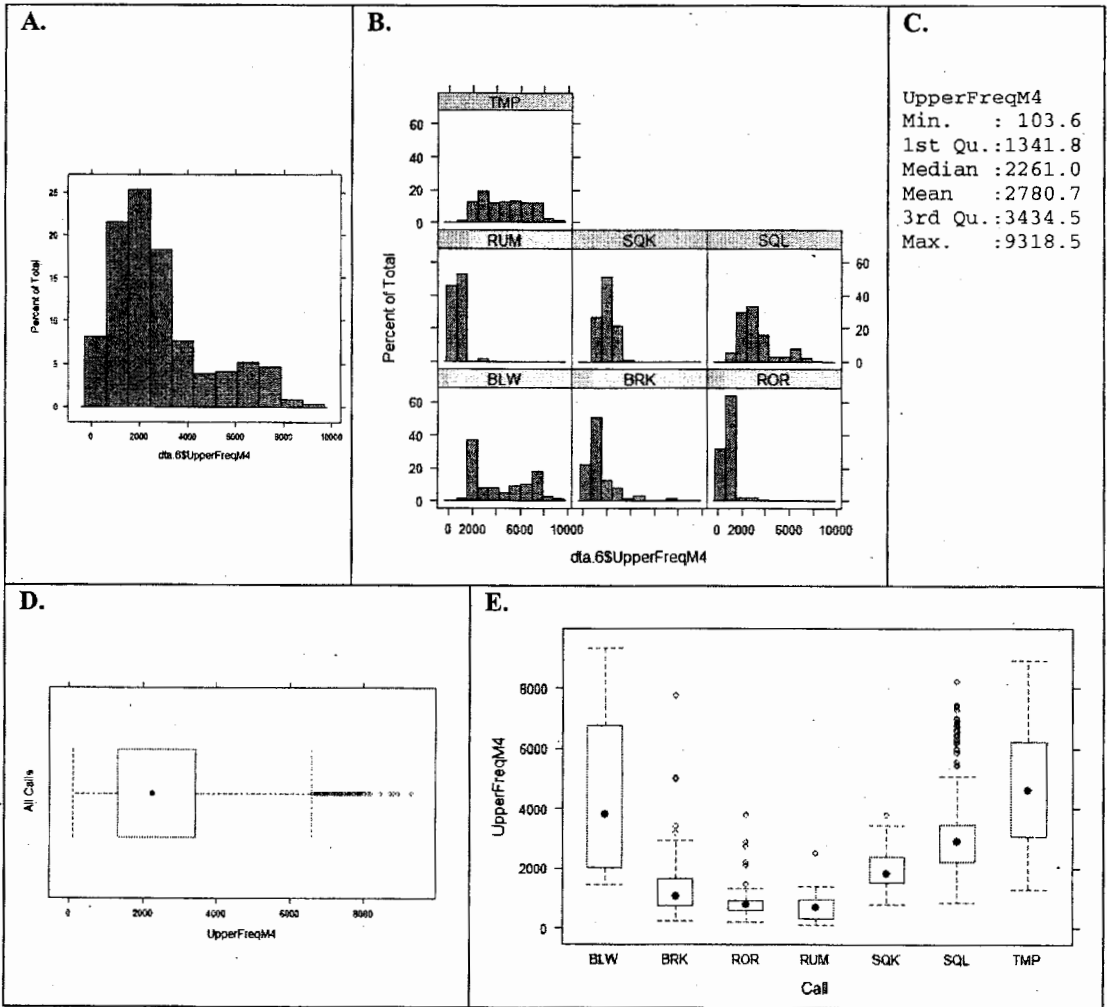
All of the histograms and boxplots below include outliers. The summary statistics exclude outliers removed. Plots are included only for the 18 parameters that were used in the statistical analysis.

The figure panels are as follows: A) Histogram of parameter for all call types combined. B) Histogram of parameter for each call type. C) Summary statistics for all call types combined. D) Box plot of parameter for all call types combined. E) Boxplot of parameter for each call type.

Abbreviations for the call types are as follow: TMP (Trumpet), RUM (Rumble), SQK (Squeak), SQL (Squeal), BLW (Blow), BRK (Bark), ROR (Roar).

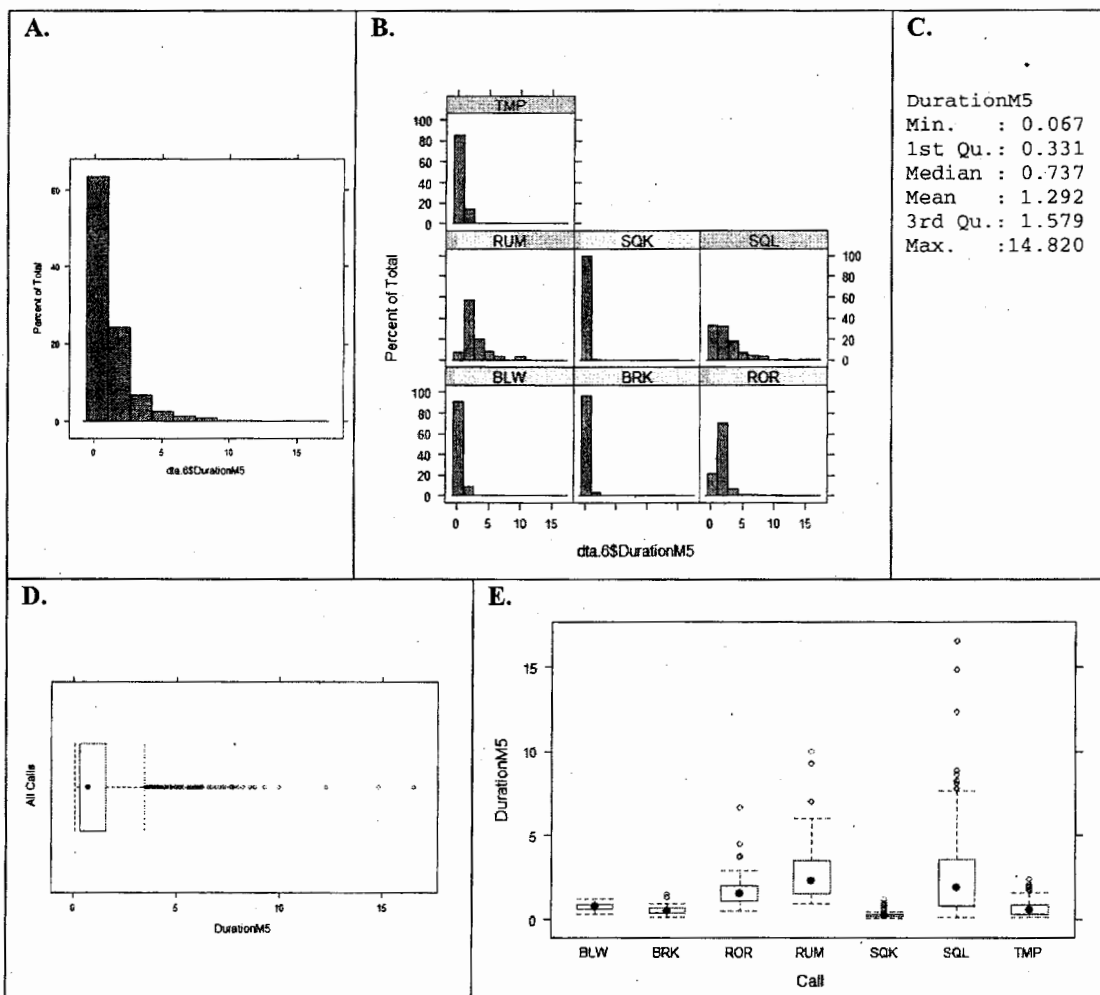


**Figure 17: Histograms and boxplots of Lower Frequency (M3)**  
 TMP and SQL more variable, other call types more stereotypic.

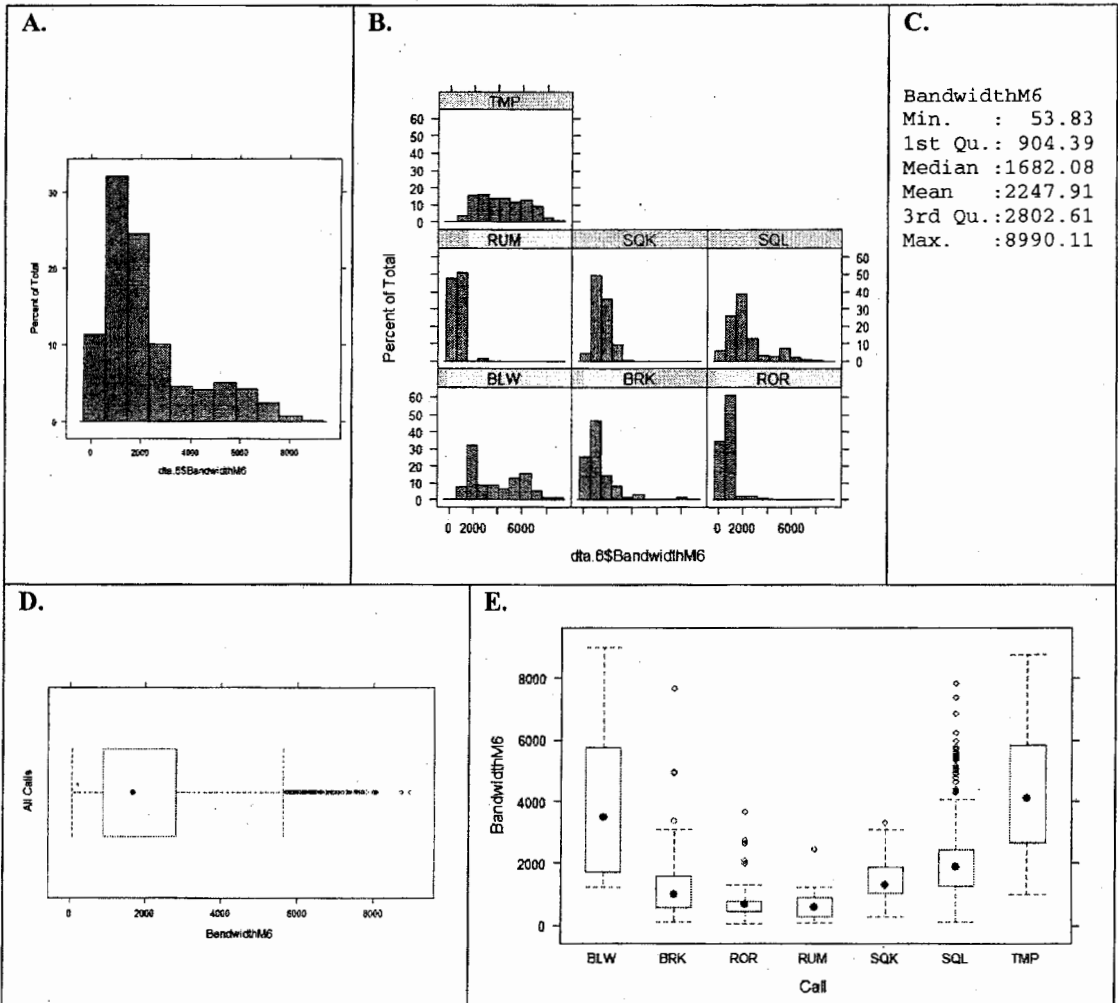


**Figure 18: Histograms and boxplots of Upper Frequency (M4)**

ROR and RUM have narrow range. TMP, SQL, and BLW highly variable, which may be partially a function of signal strength (distance to caller).

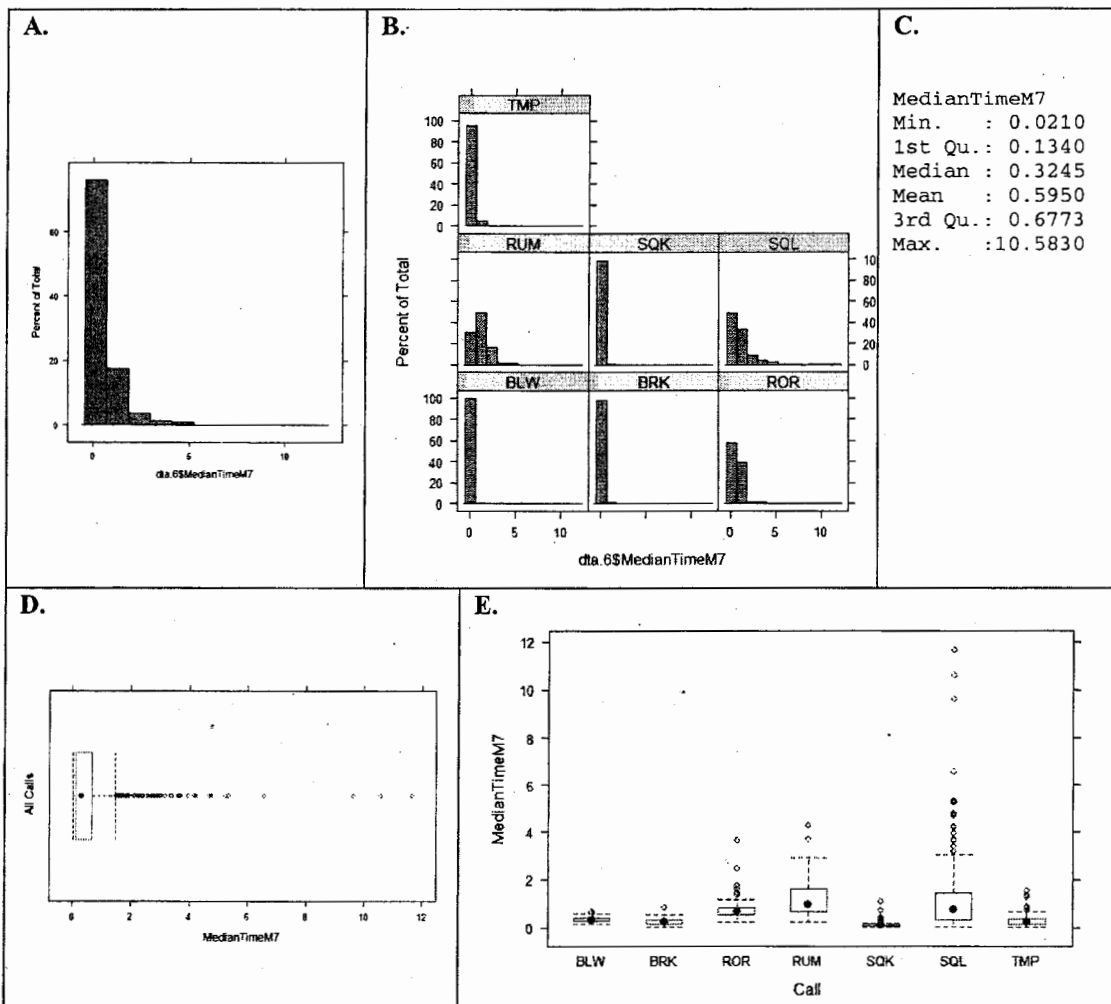


**Figure 19: Histograms and boxplots of Duration (M5)**  
 ROR and RUM overlap in perception. SQL variable in duration. SQK, BRK, and TMP stereotypic.

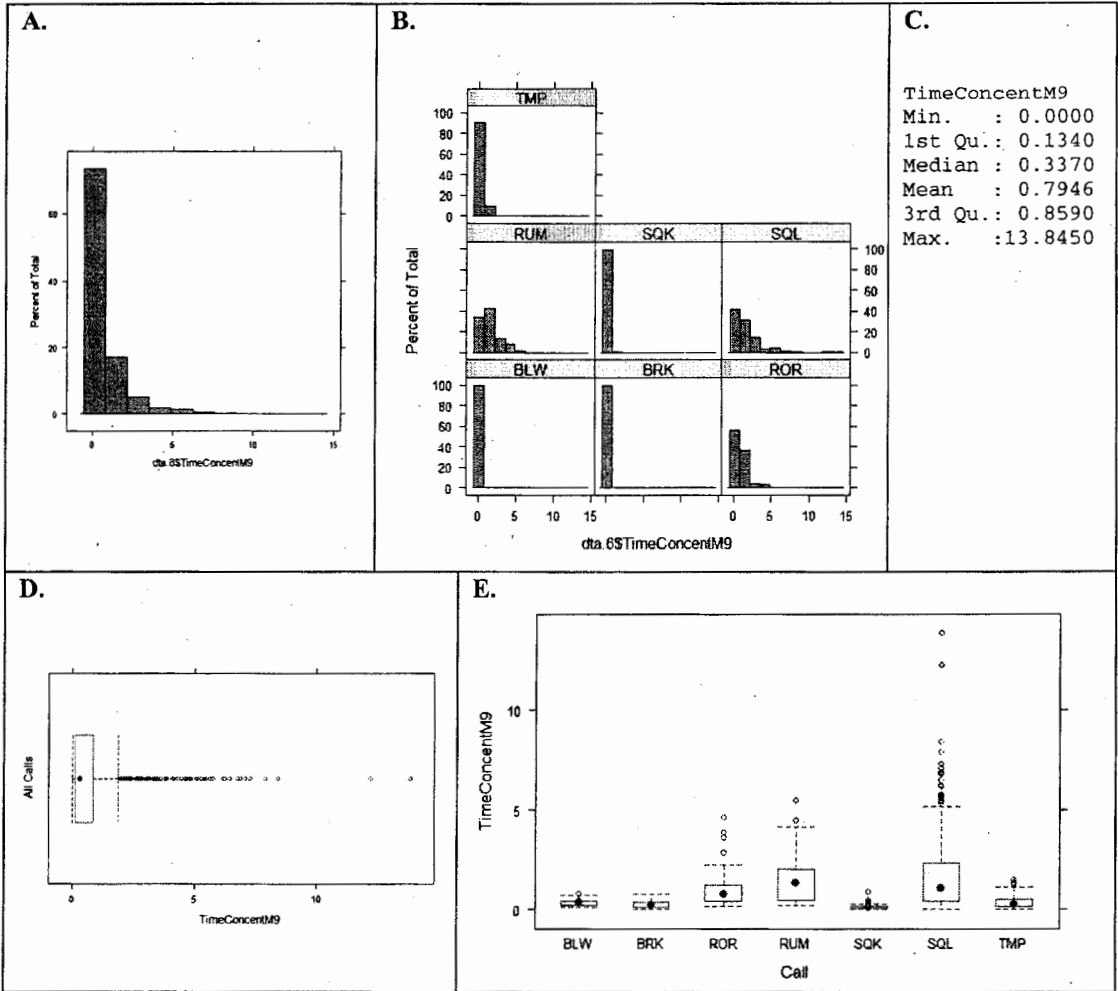


**Figure 20: Histograms and boxplots of Bandwidth (M6)**

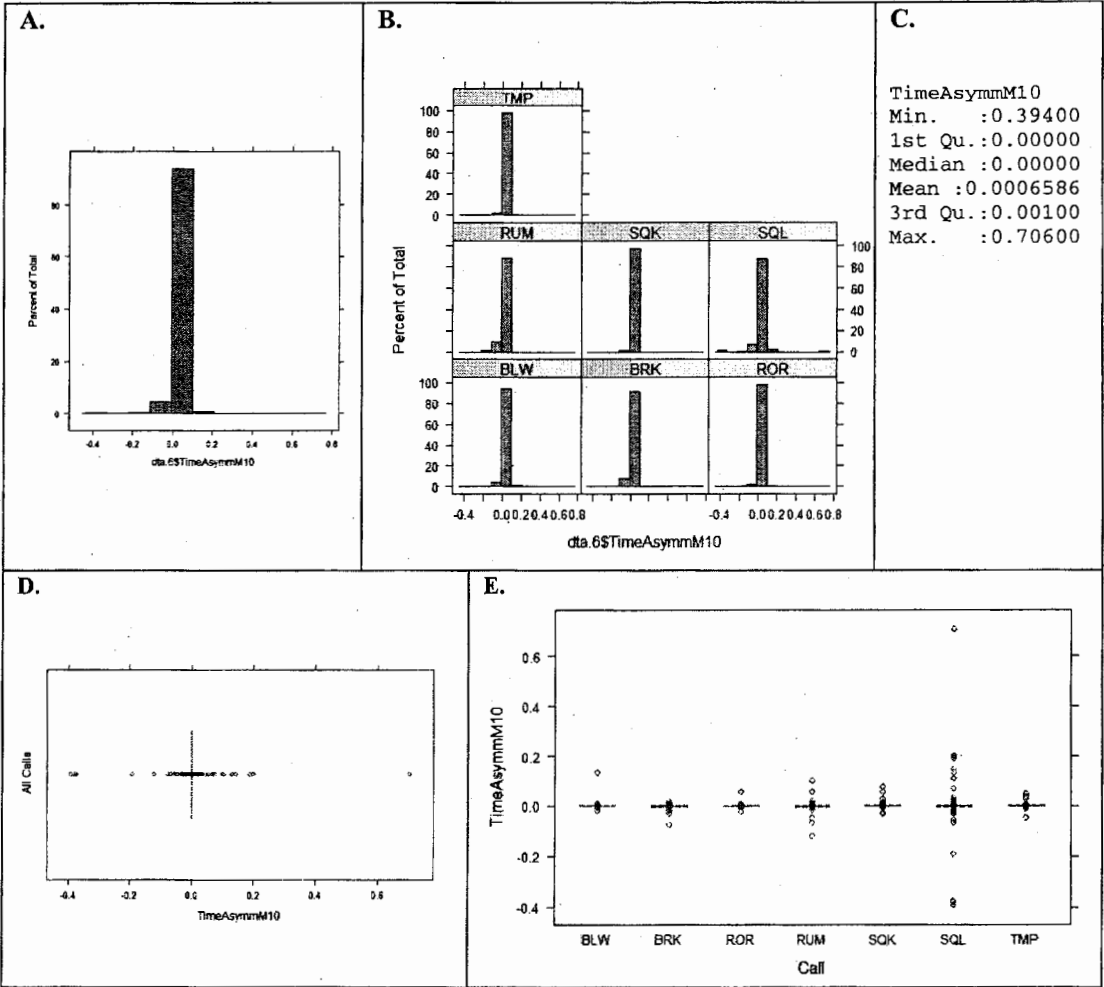
RUM mostly low frequency (as expected). BLW and TMP have very high bandwidth and variability, which may partially be due to distance to source. Squeal has relatively high bandwidth and very high variability. SQK has relatively wide bandwidth, but it is stereotypic.



**Figure 21: Histograms and boxplots of Median Time (M7)**



**Figure 22: Histograms and boxplots of Temporal Concentration (M9) SQL highly variable. RUM more variable than others.**



**Figure 23: Histograms and boxplots of Temporal Asymmetry (M10)**  
 Stereotypic with minimal variability.



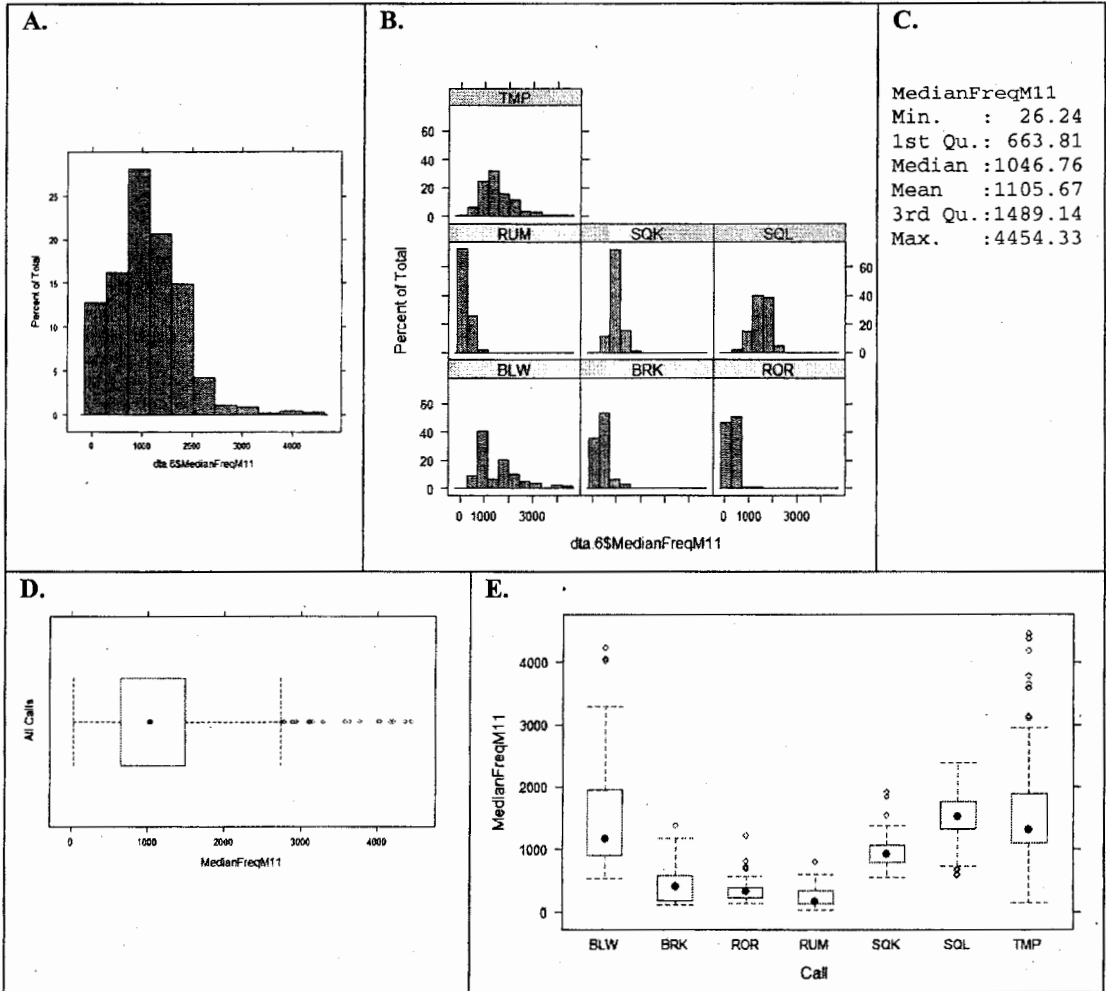


Figure 24: Histograms and boxplots of Median Frequency (M11)

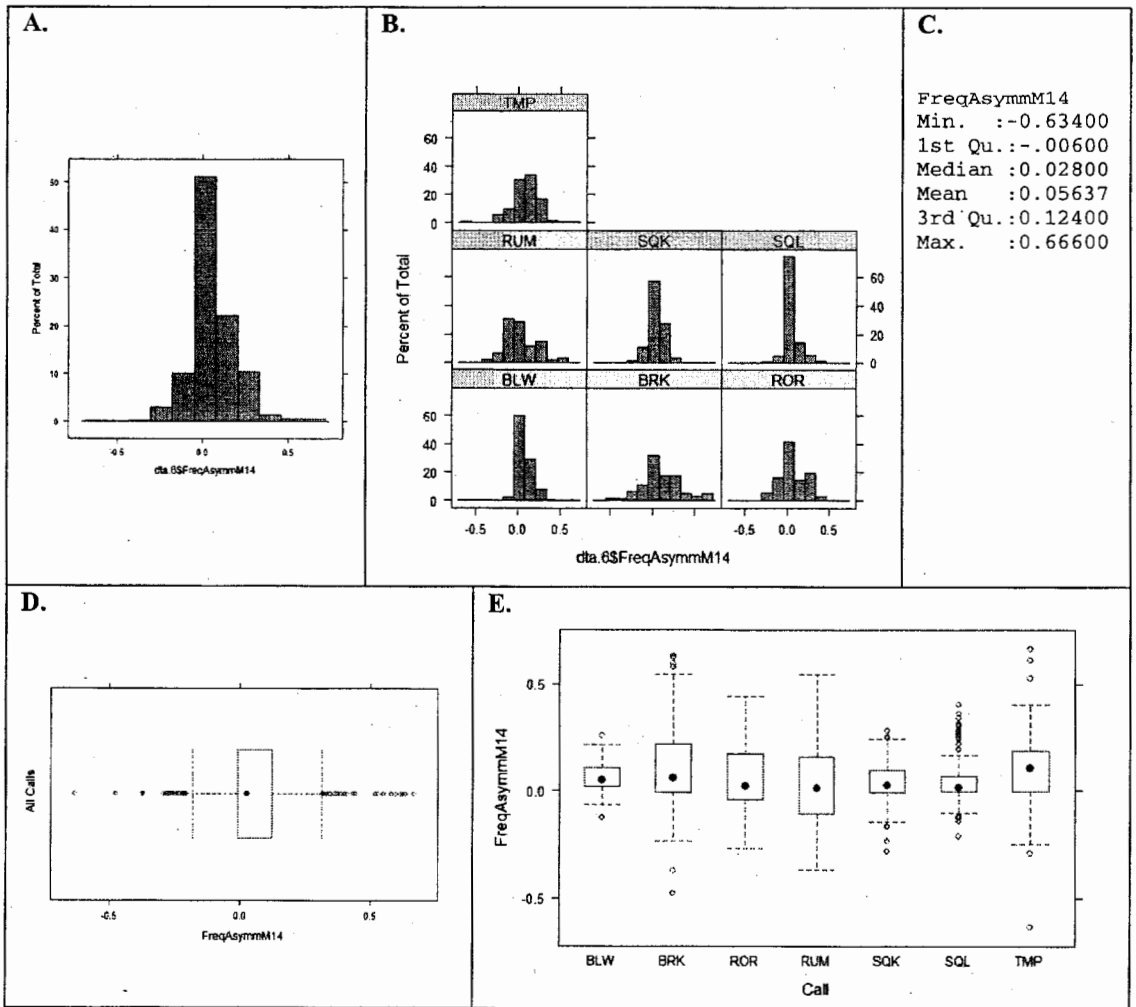
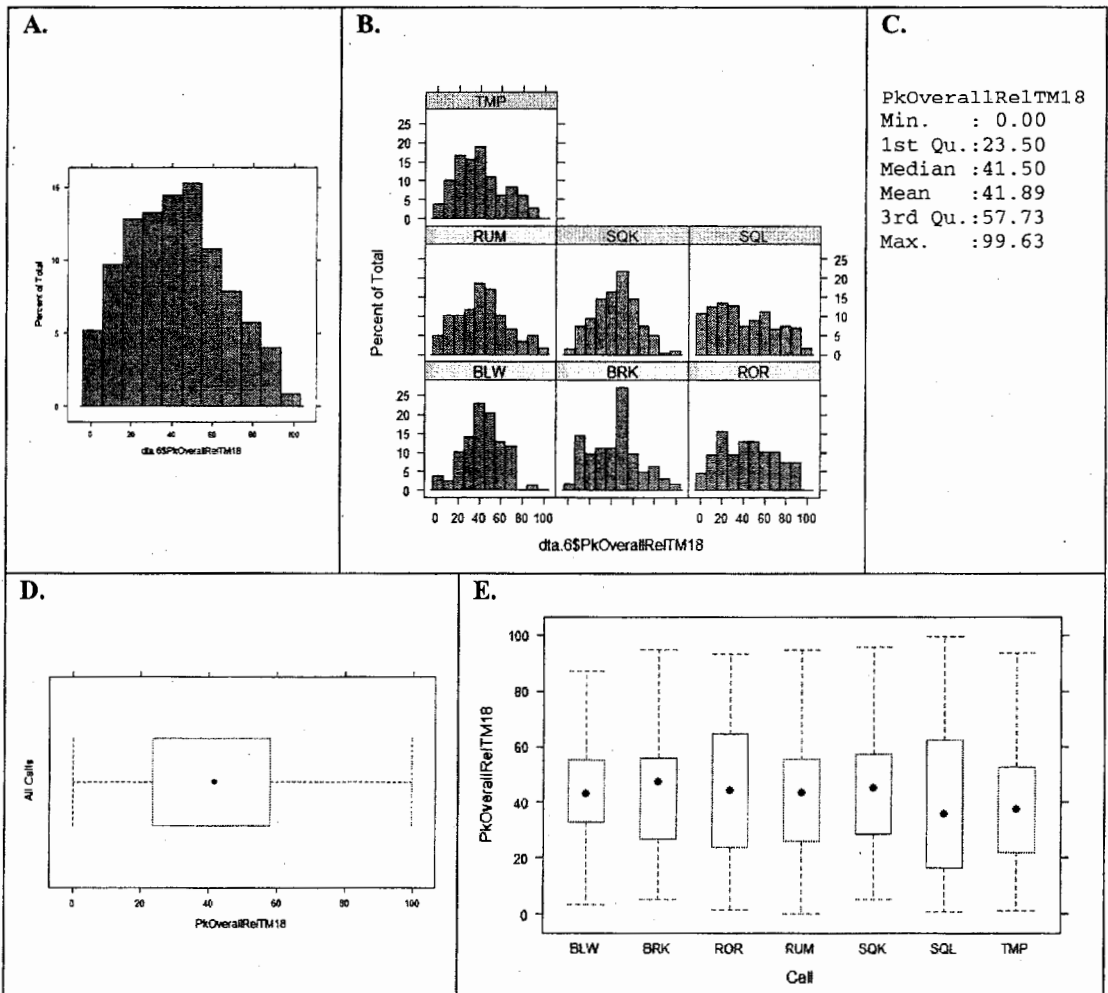
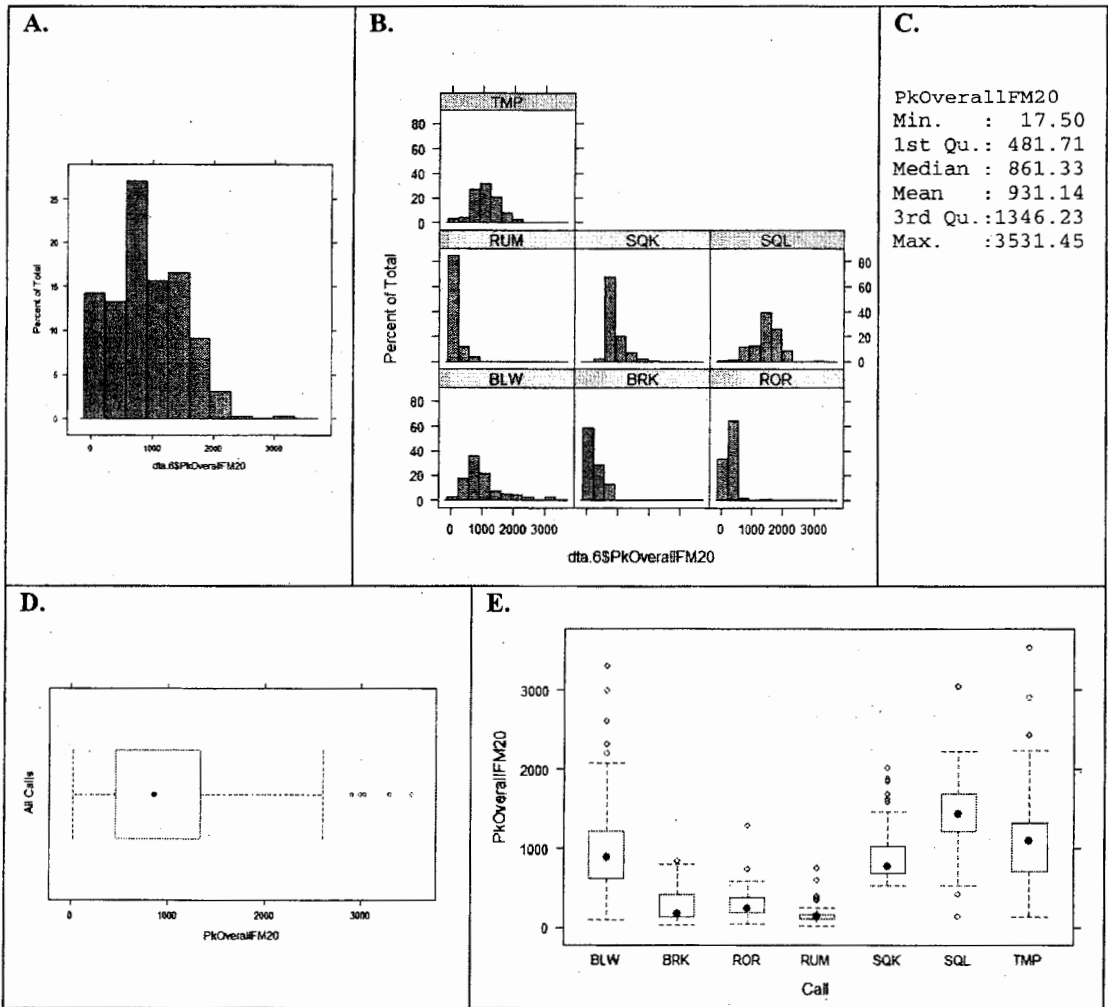


Figure 25: Histograms and boxplots of Frequency Asymmetry (M14)



**Figure 26: Histograms and boxplots of Relative Time of Peak Overall Intensity (M18)**  
 High variability within call type, but similar among call types.



**Figure 27: Histograms and boxplots of Frequency of Peak Overall Intensity (M20)**

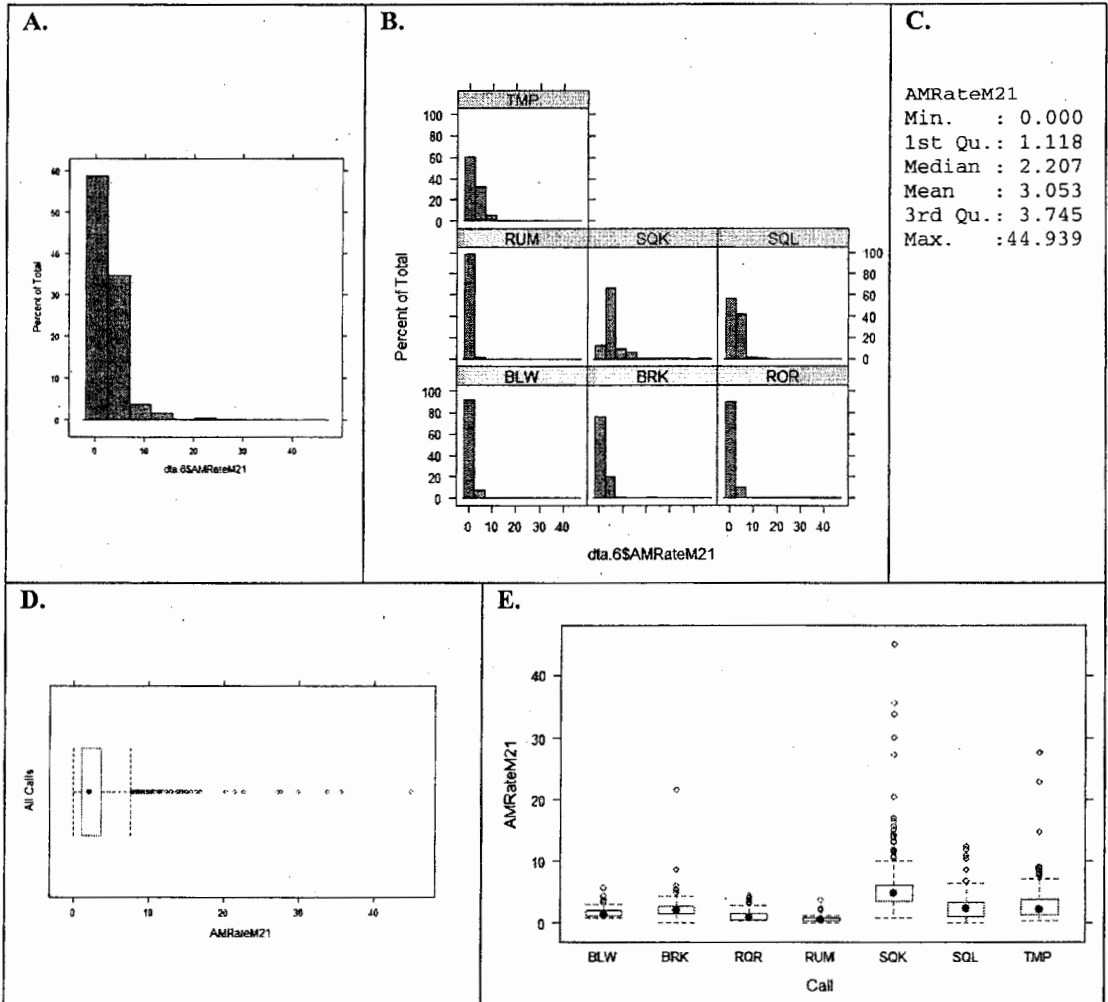


Figure 28: Histograms and boxplots of AM Rate (M21)

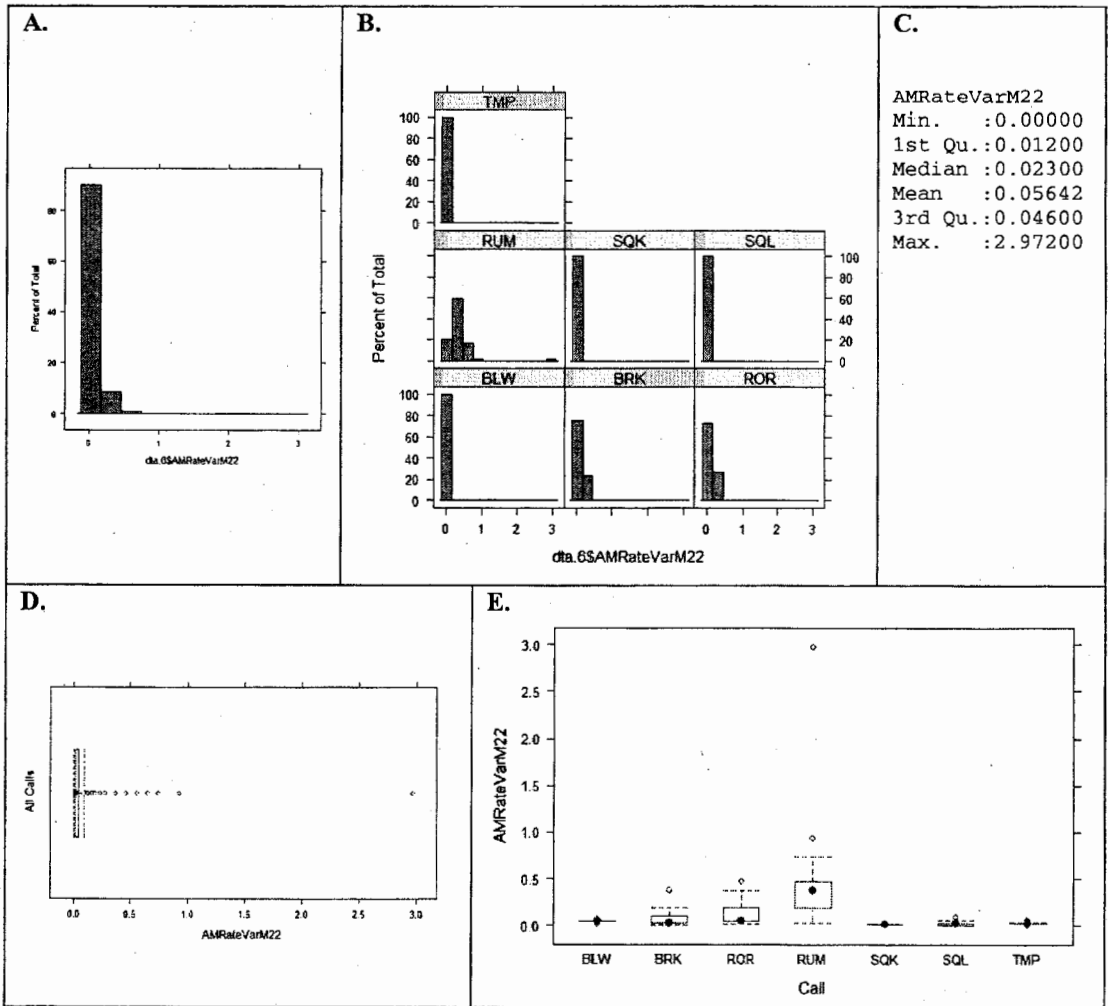


Figure 29: Histograms and boxplots of AM Rate Variation (M22)

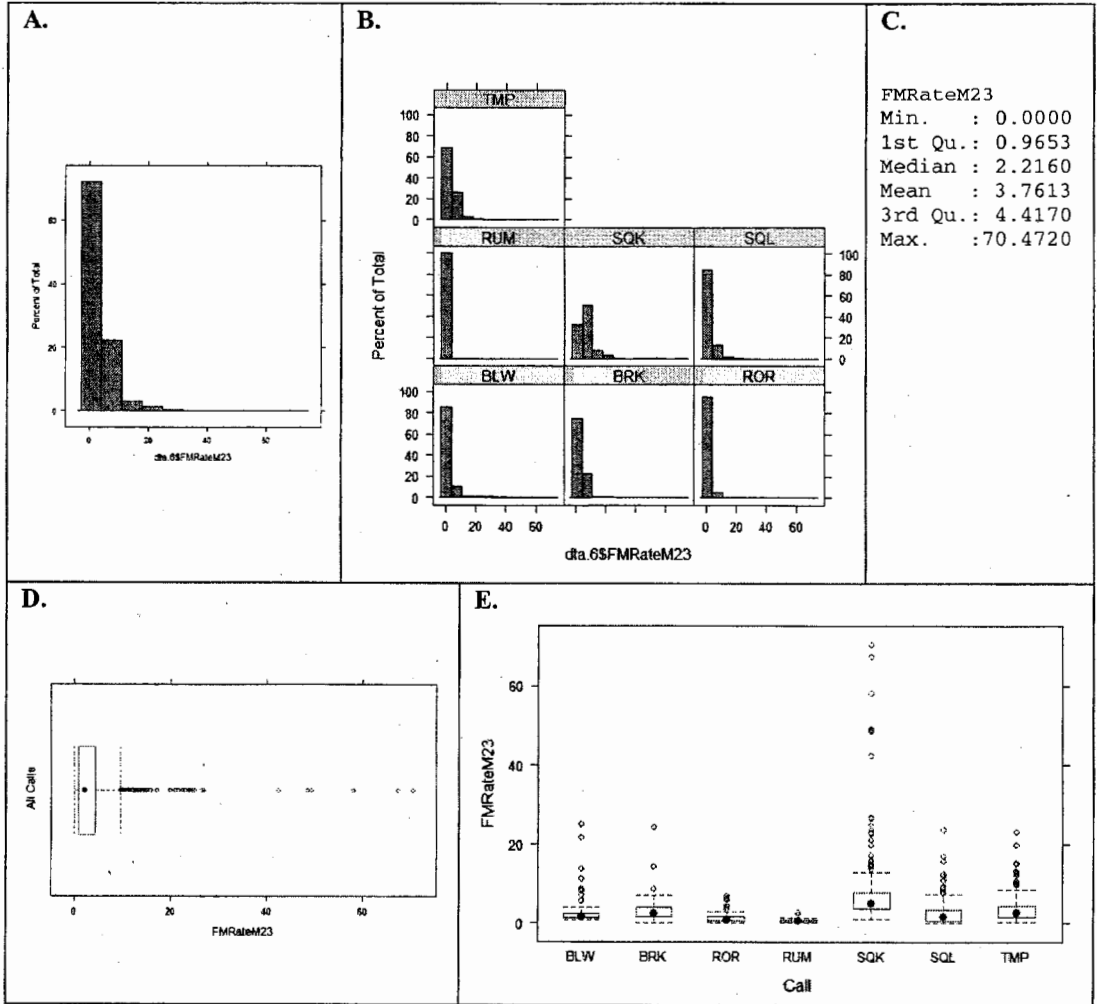


Figure 30: Histograms and boxplots of FM Rate (M23)

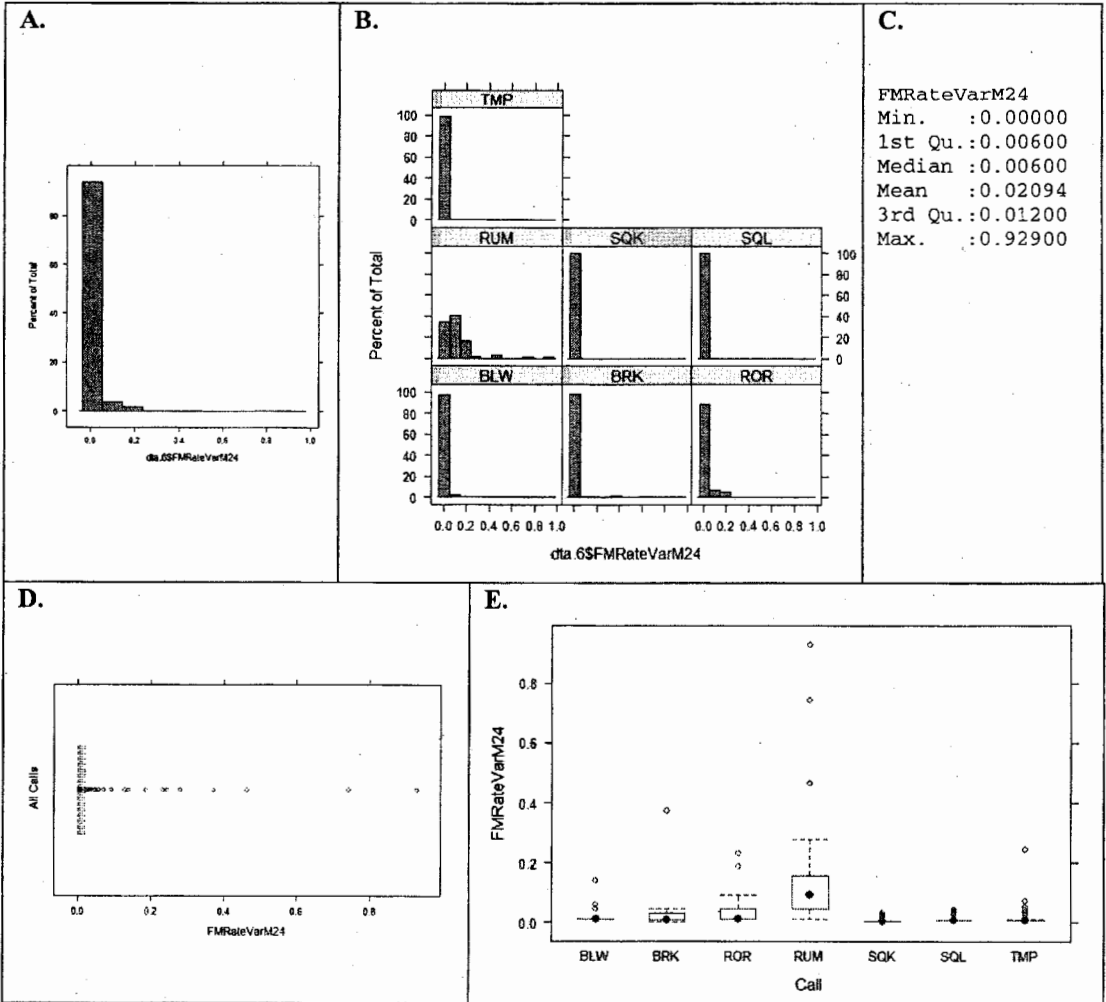
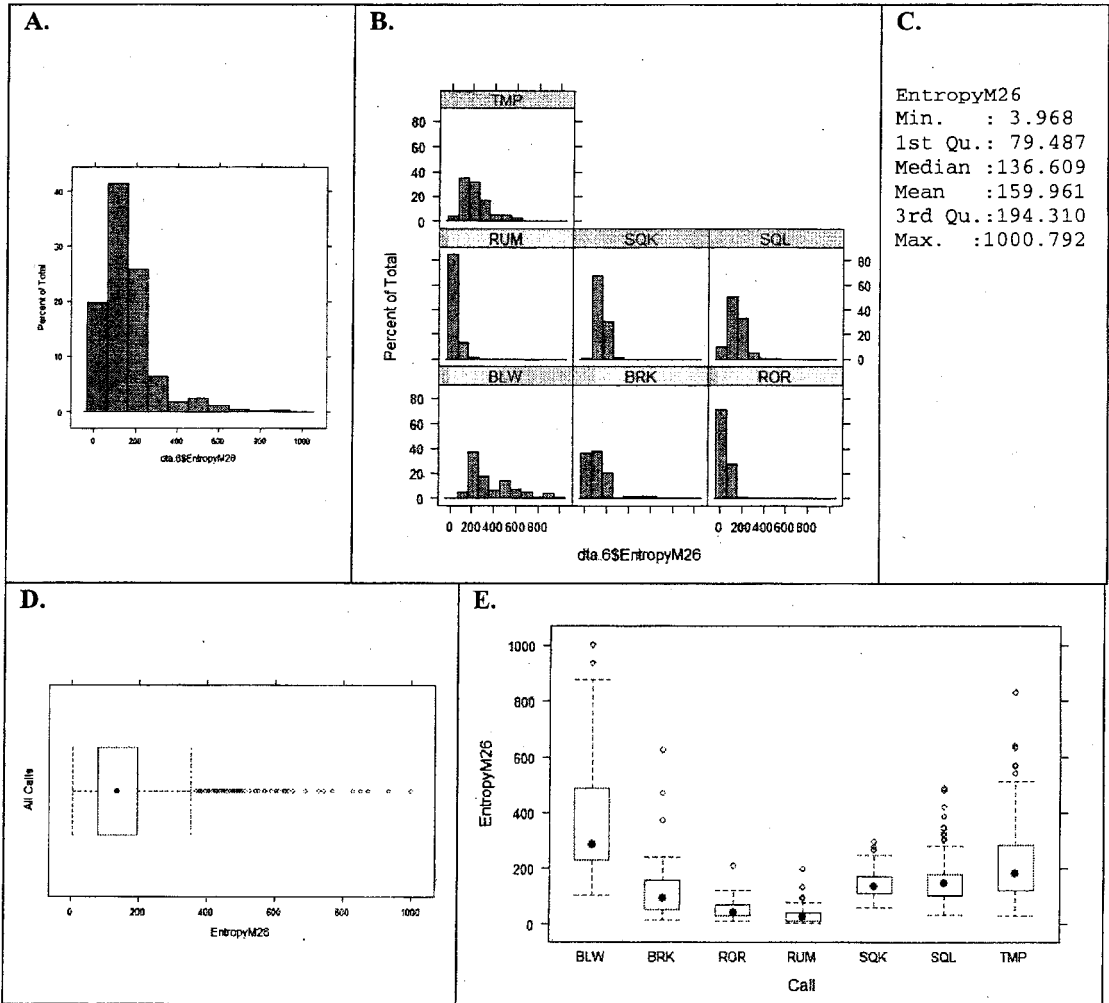


Figure 31: Histograms and boxplots of FM Rate Variation (M24)





**Figure 32: Histograms and boxplots of Overall Entropy (M26)**

Highly variable among call types. TMP has highest variability (tonal-noisy variability), which may reflect level of arousal in switch to noise. Wider bandwidth calls have higher entropy. Chaotic noise has been perceived in the calls with higher frequency ranges (roar, trumpet, squeals), so this is consistent with perceptual aural cues.

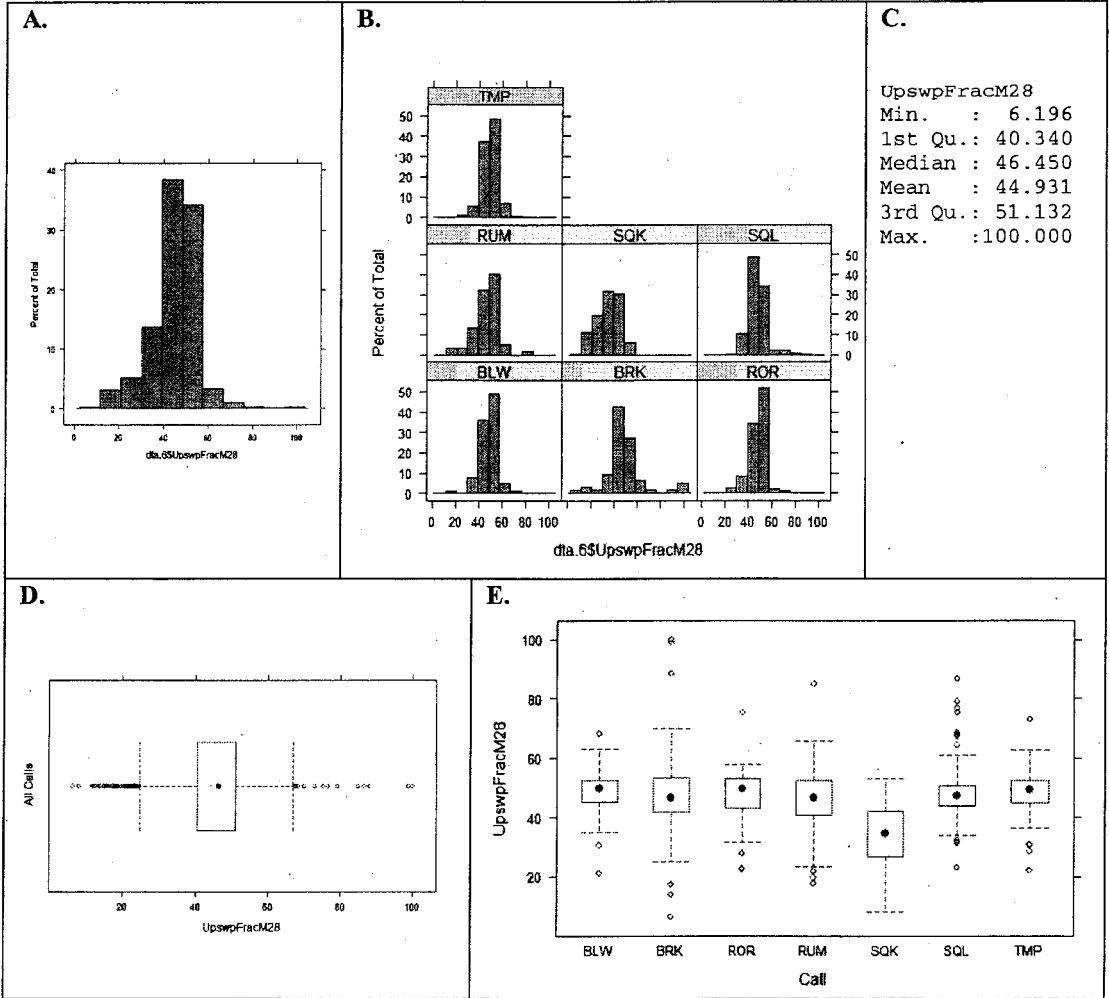


Figure 33: Histograms and boxplots of Upsweep Fraction (M28)

### Q-Q PLOTS OF SCALED DATA

In a Q-Q plot, the x-axis is the expected value for a normal distribution and the y-axis is observed. The data are ranked and the quartile is calculated. If data are distributed normally then the shape will be a straight line where  $y=x$ .

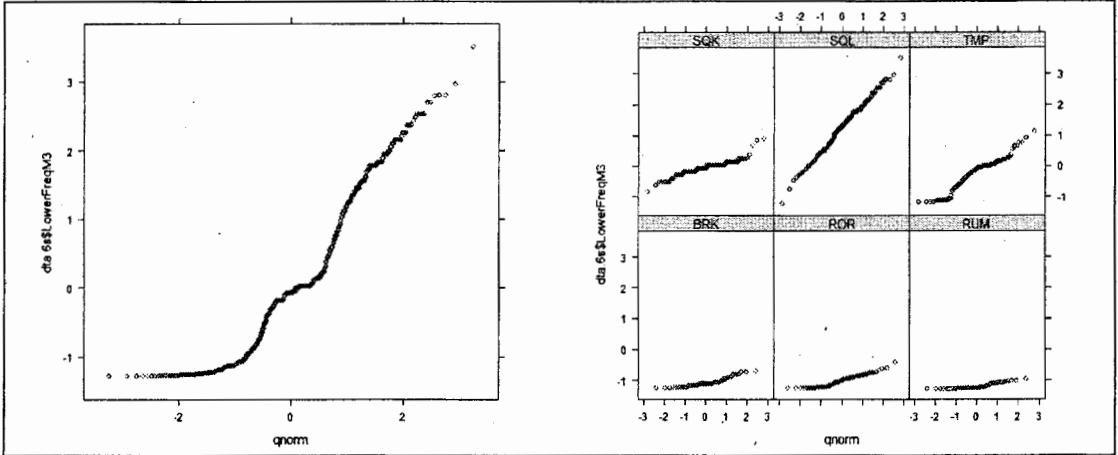


Figure 34: Q-Q plot of Lower Frequency (M3)

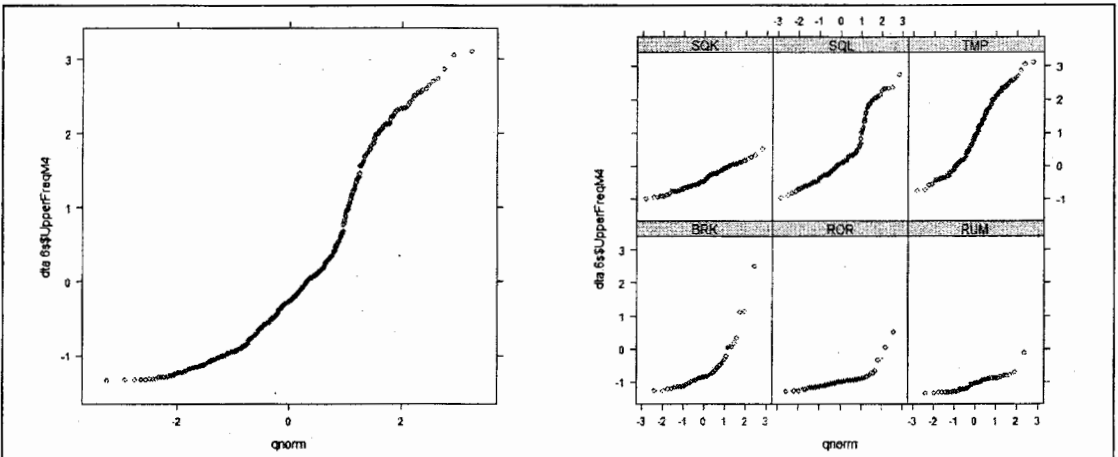


Figure 35: Q-Q plots of Upper Frequency (M4)

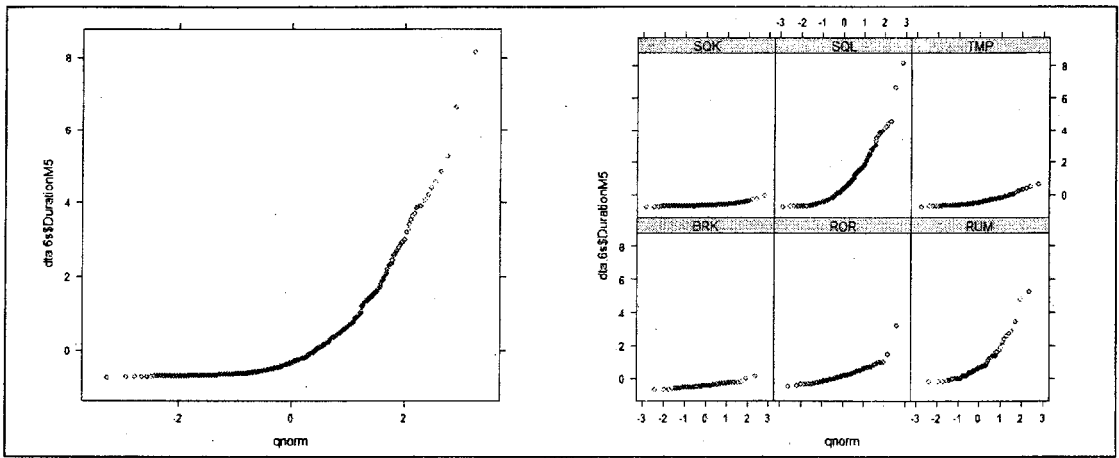


Figure 36: Q-Q plots of Duration (M5)

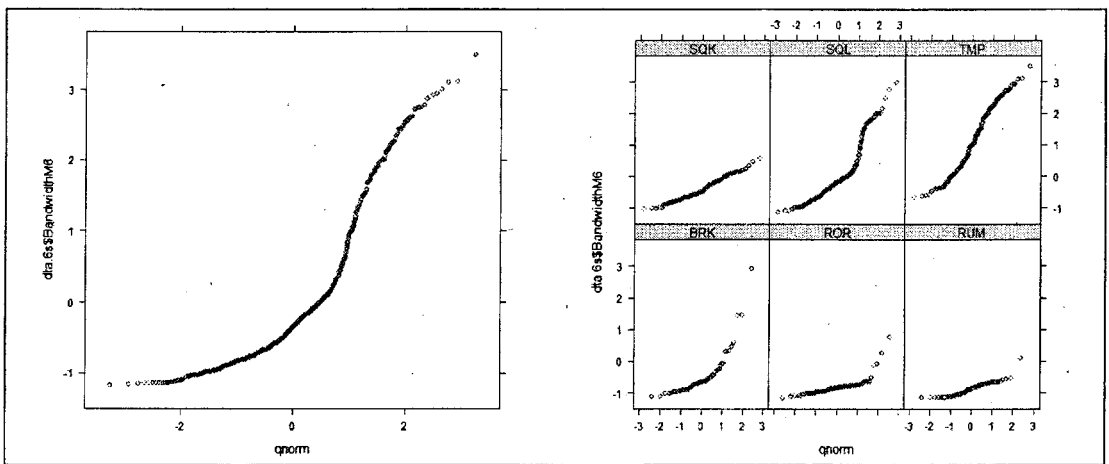


Figure 37: Q-Q plots of Bandwidth (M6)

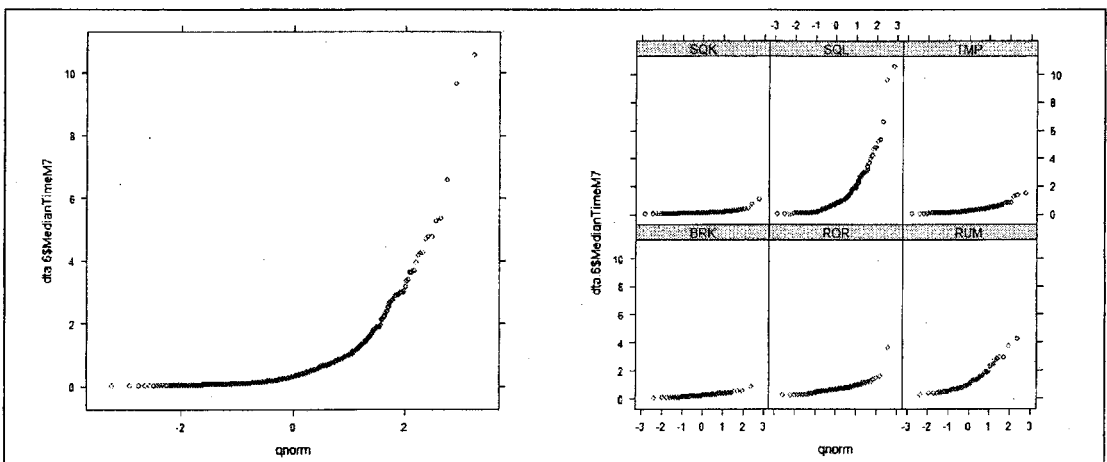


Figure 38: Q-Q plots of Median Time (M7)

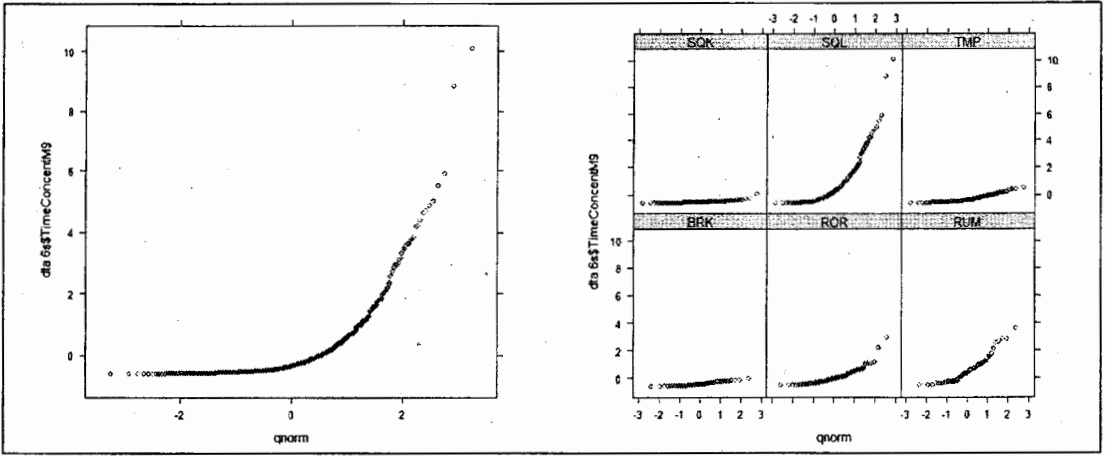


Figure 39: Q-Q plots of Temporal Concentration (M9)

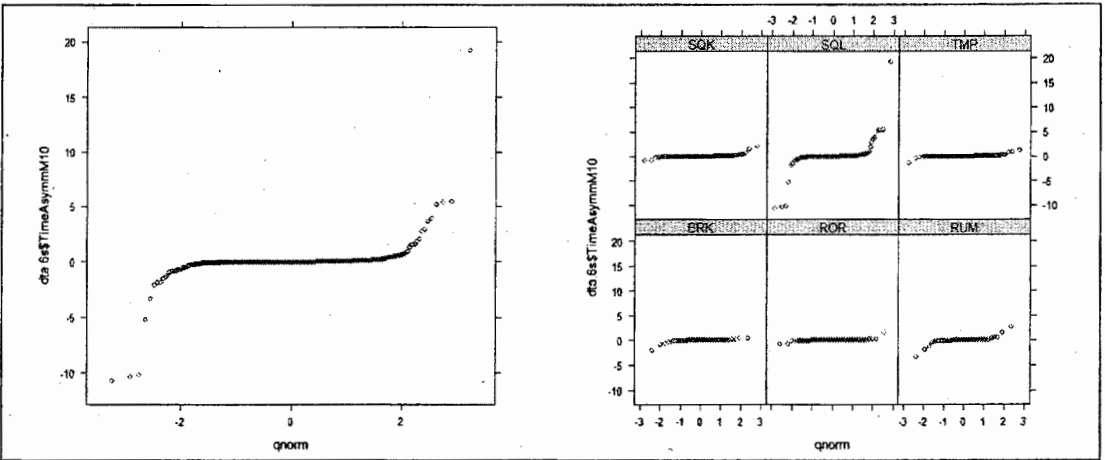


Figure 40: Q-Q plots of Temporal Asymmetry (M10)

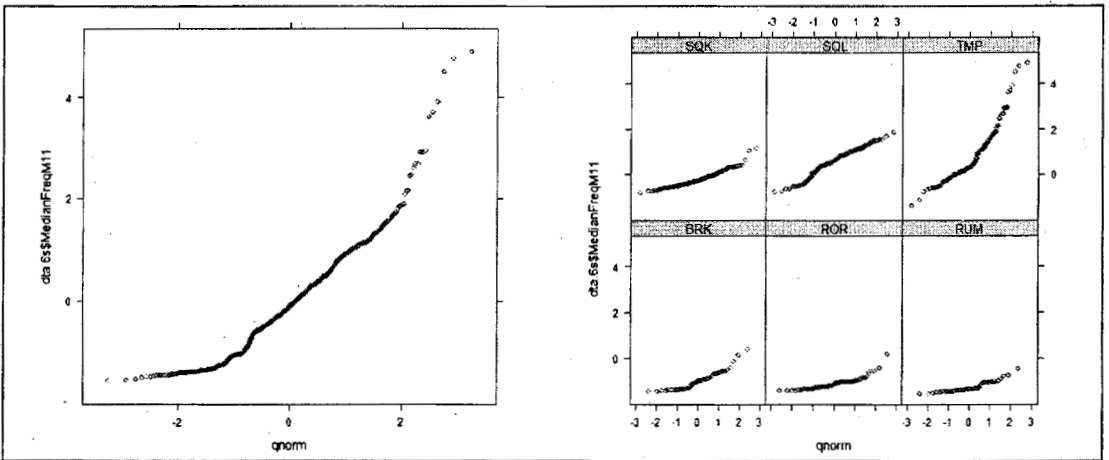


Figure 41: Q-Q plots of Median Frequency (M11)

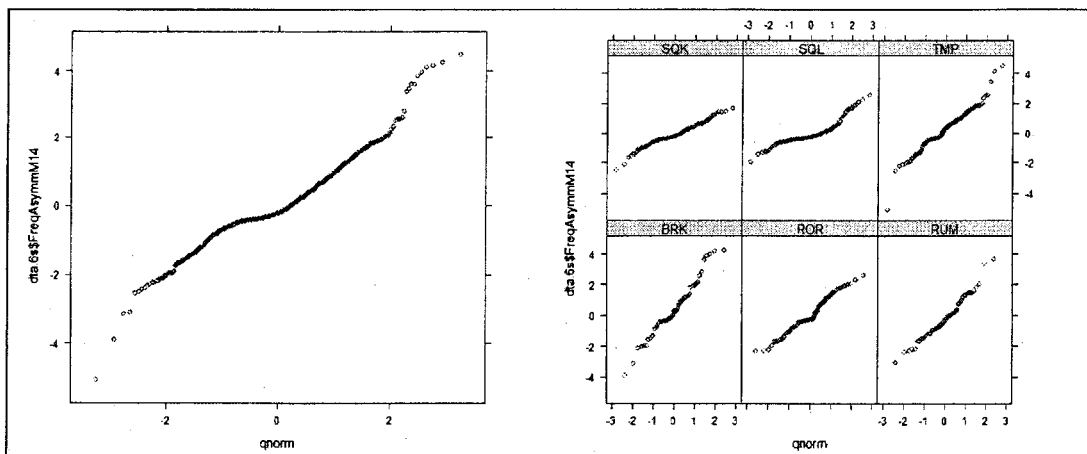


Figure 42: Q-Q plots of Frequency Asymmetry (M14)

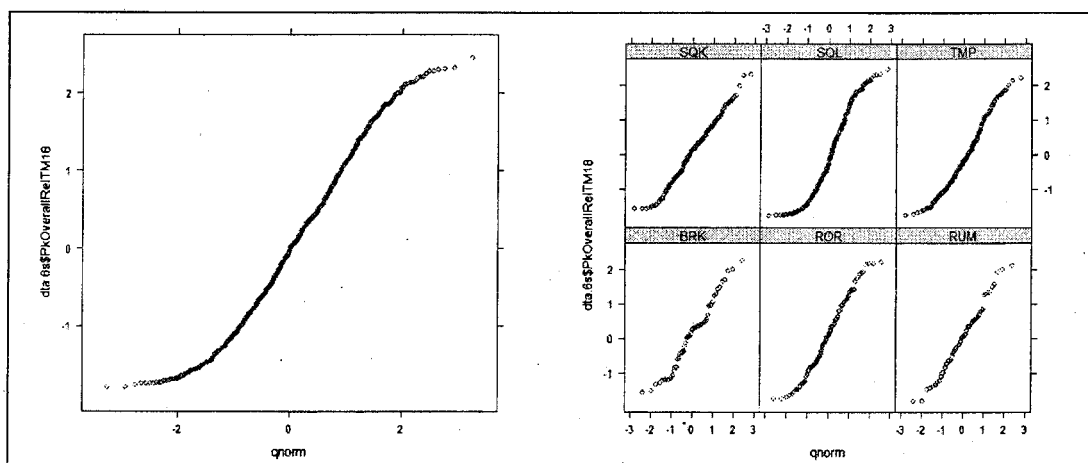


Figure 43: Q-Q plots of Relative Time of Peak Overall Intensity (M18)

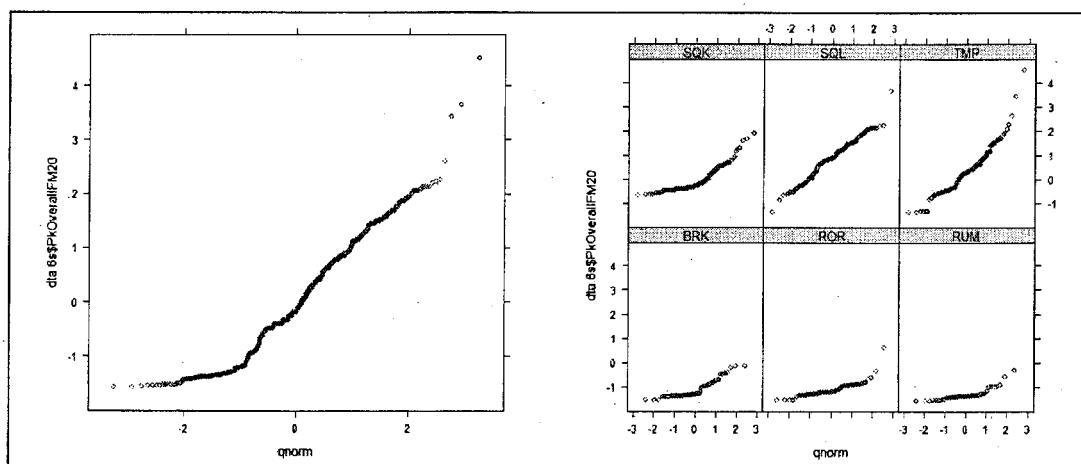


Figure 44: Q-Q plots of Frequency of Peak Overall Intensity (M20)

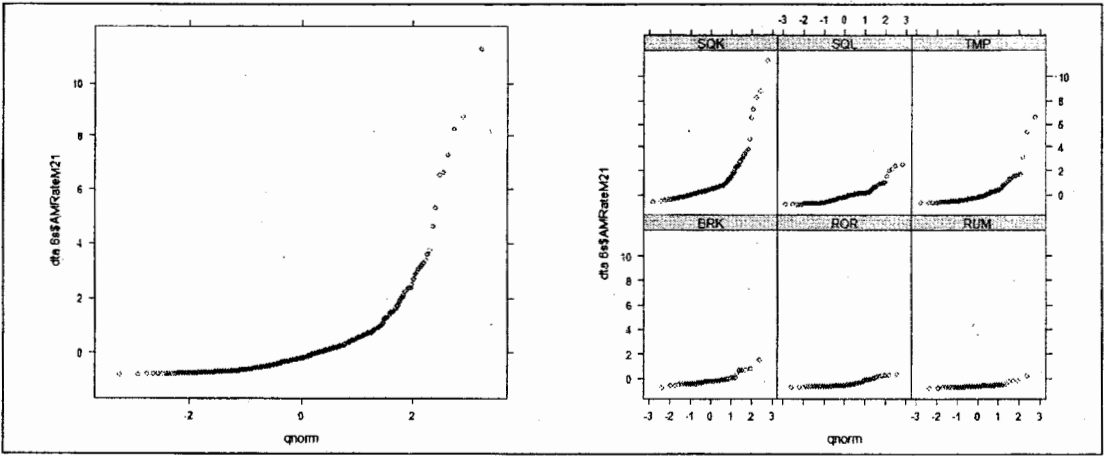


Figure 45: Q-Q plots of AM Rate (M21)

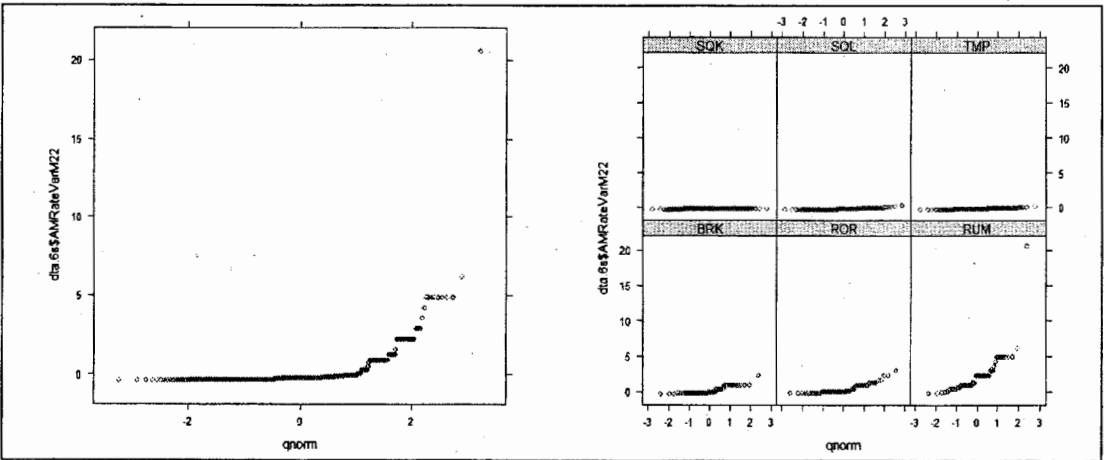


Figure 46: Q-Q plots of AM Rate Variation (M22)

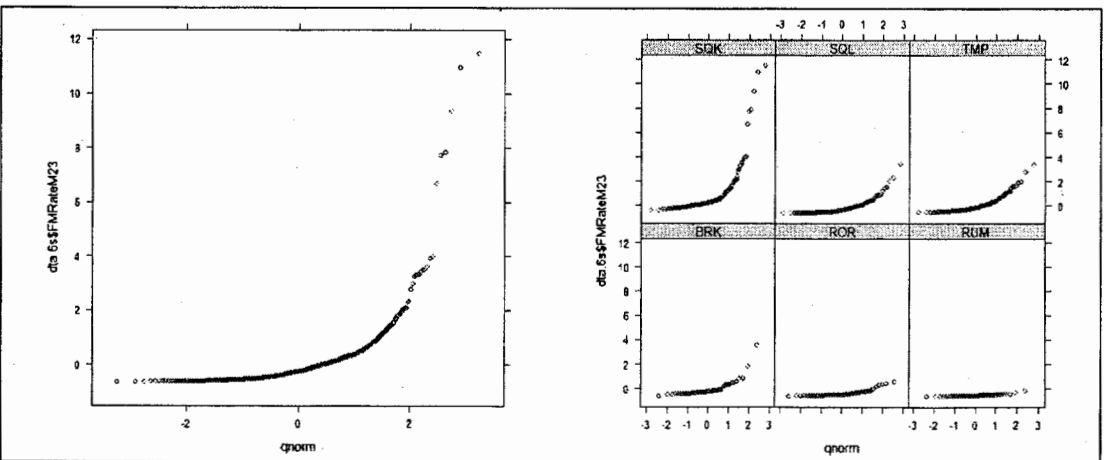


Figure 47: Q-Q plots of FM Rate (M23)

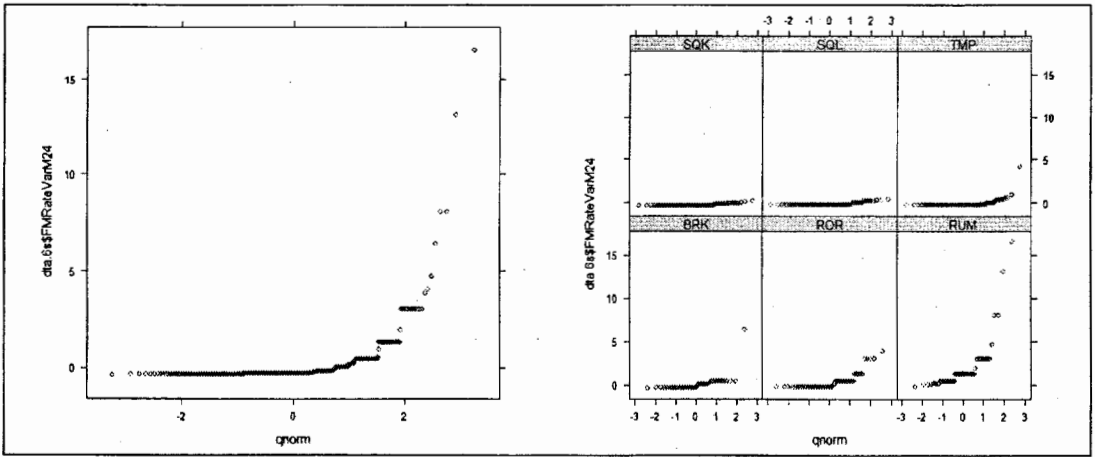


Figure 48: Q-Q plots of FM Rate Variation (M24)

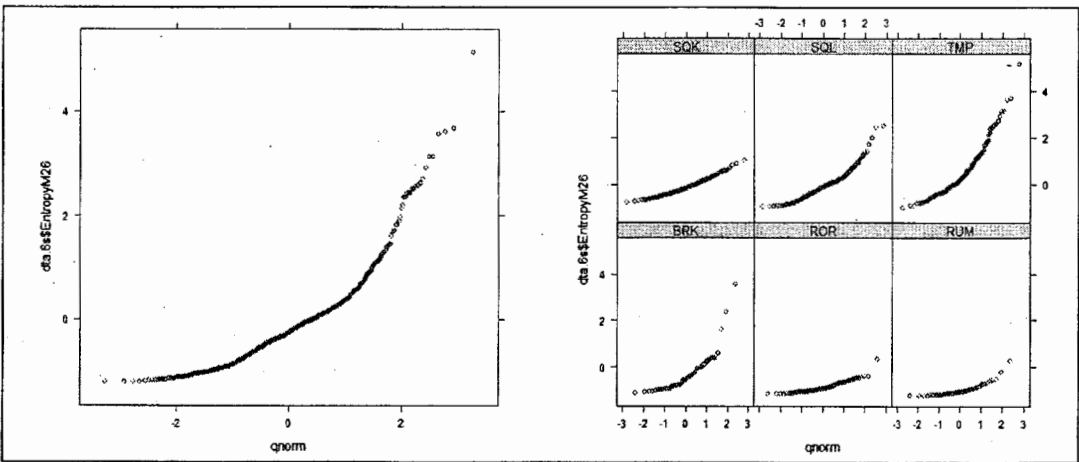


Figure 49: Q-Q plots of Overall Entropy (M26)

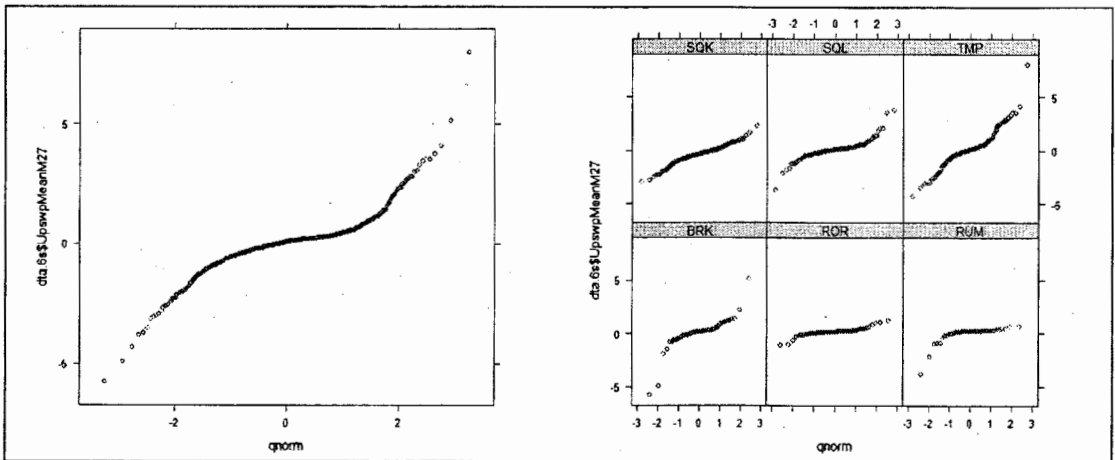


Figure 50: Q-Q plots of Upsweep Mean (M27)



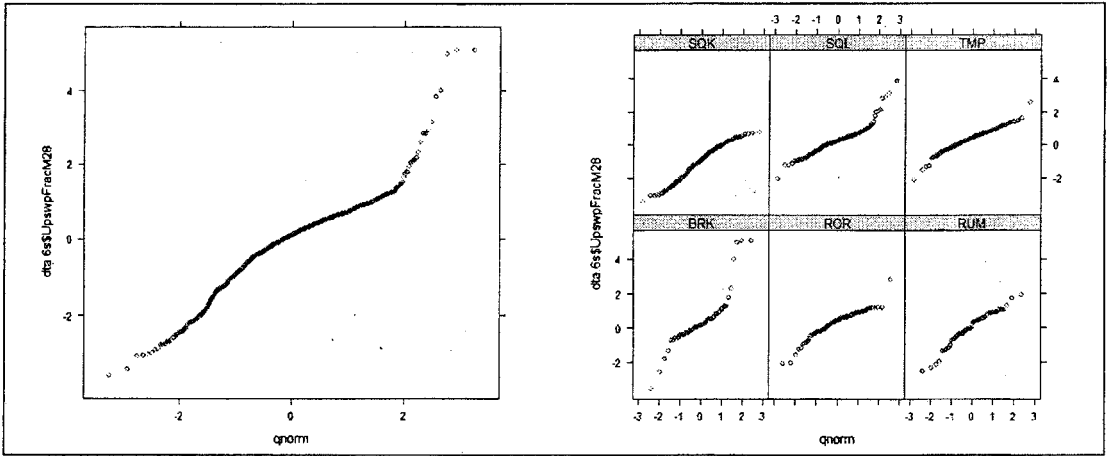
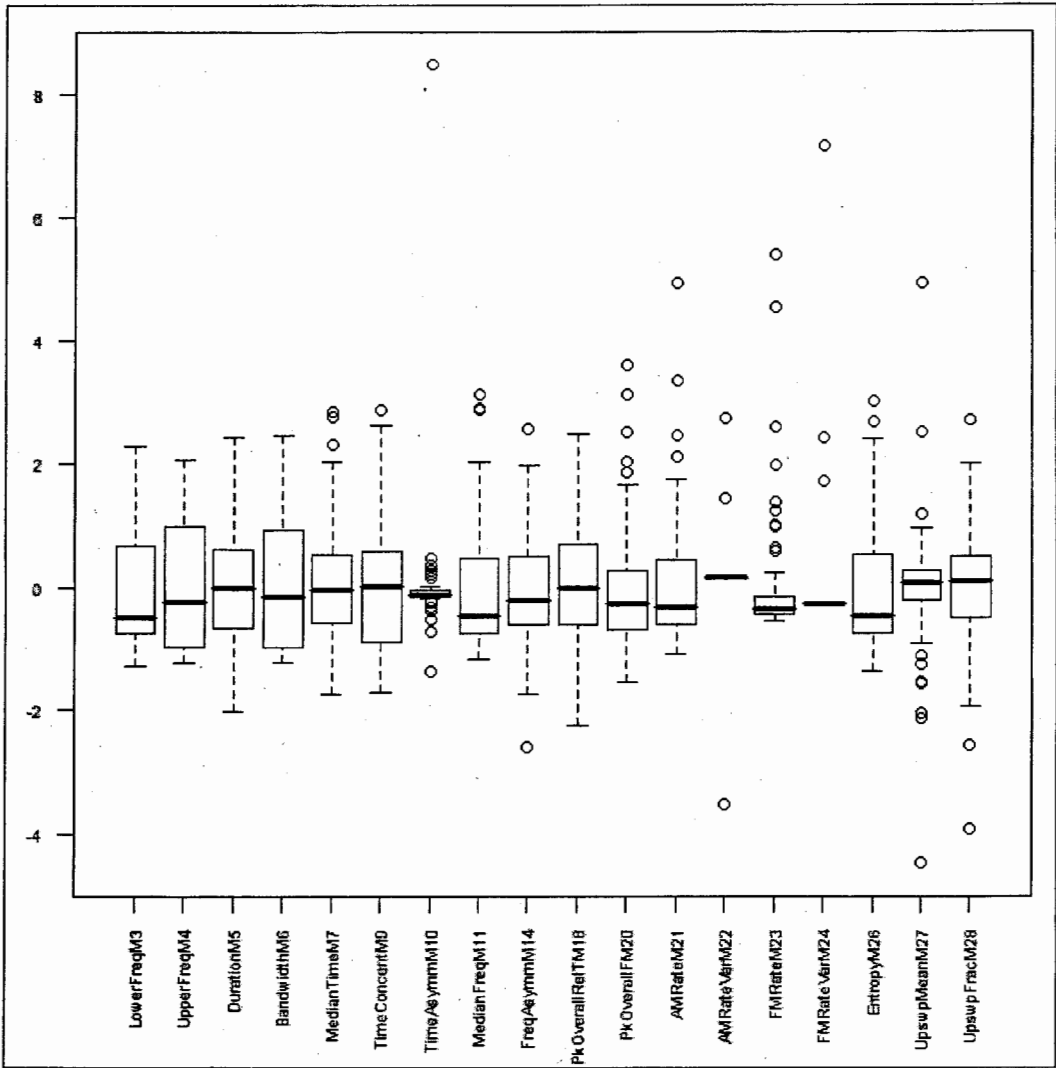
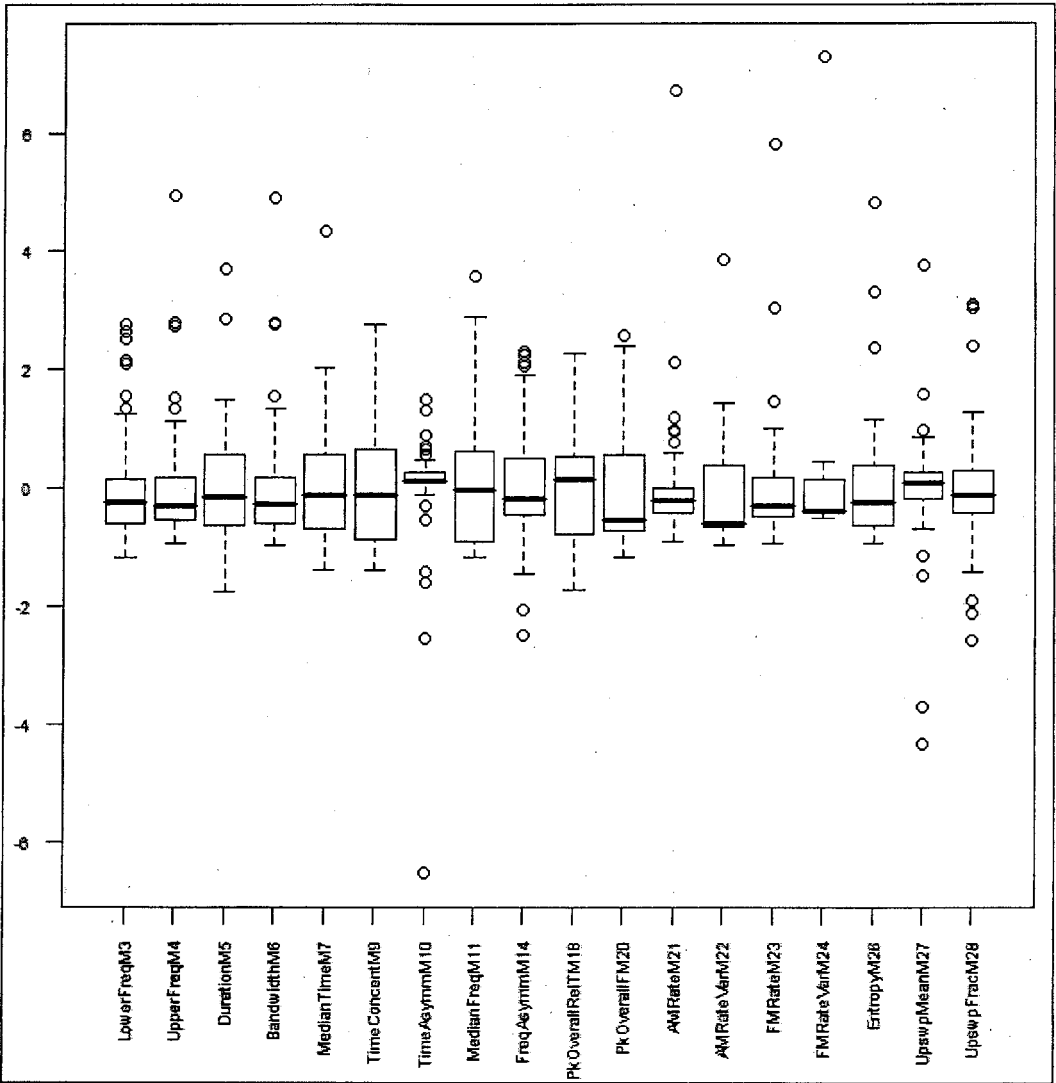


Figure 51: Q-Q plots of Upsweep Fraction (M28)

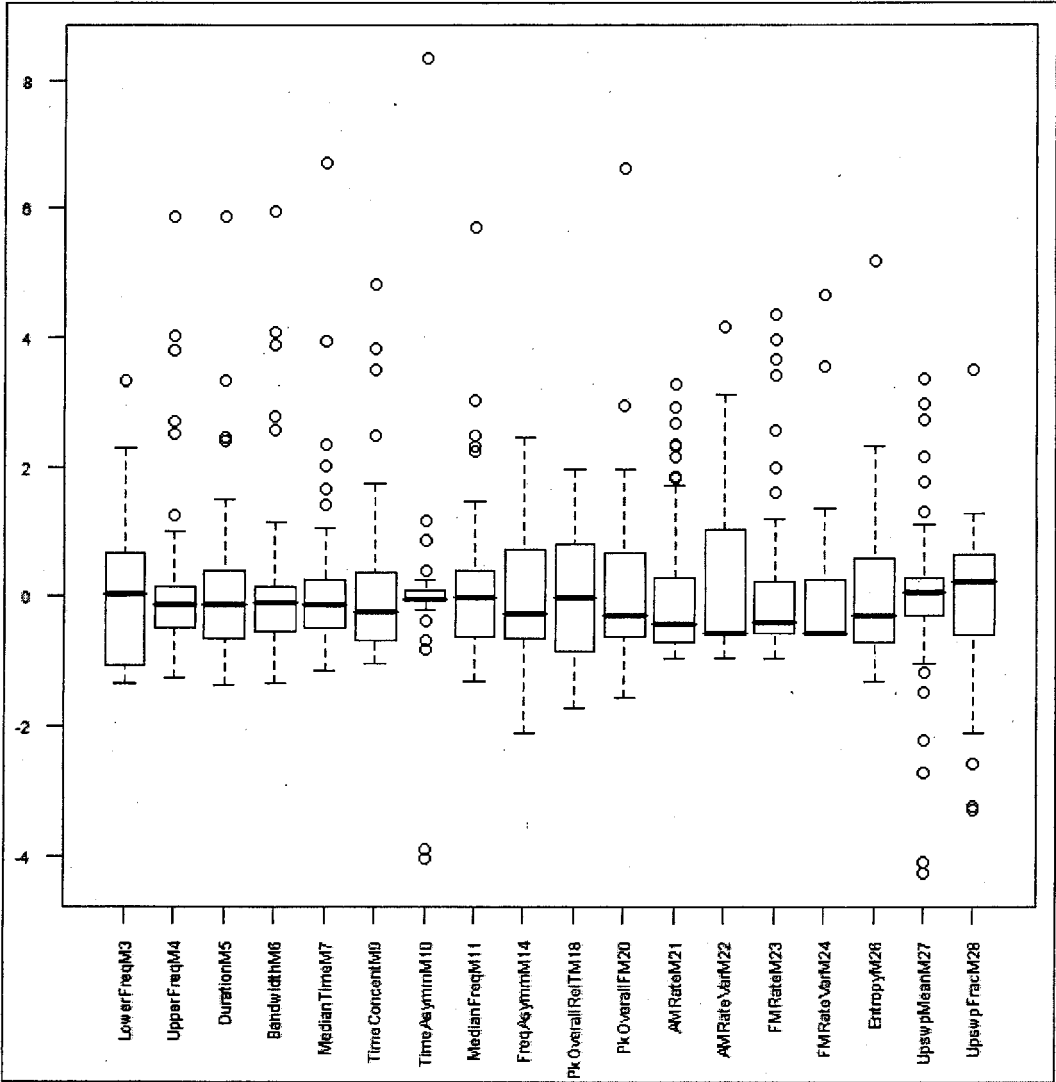
**BOX PLOTS OF SCALED VARIABLES FOR IDENTIFYING OUTLIERS**



**Figure 52: Boxplot of scaled variables for Blow**  
 Acoustic parameters are on the x-axis. Scaled (normalized) values of the parameters are on the y-axis

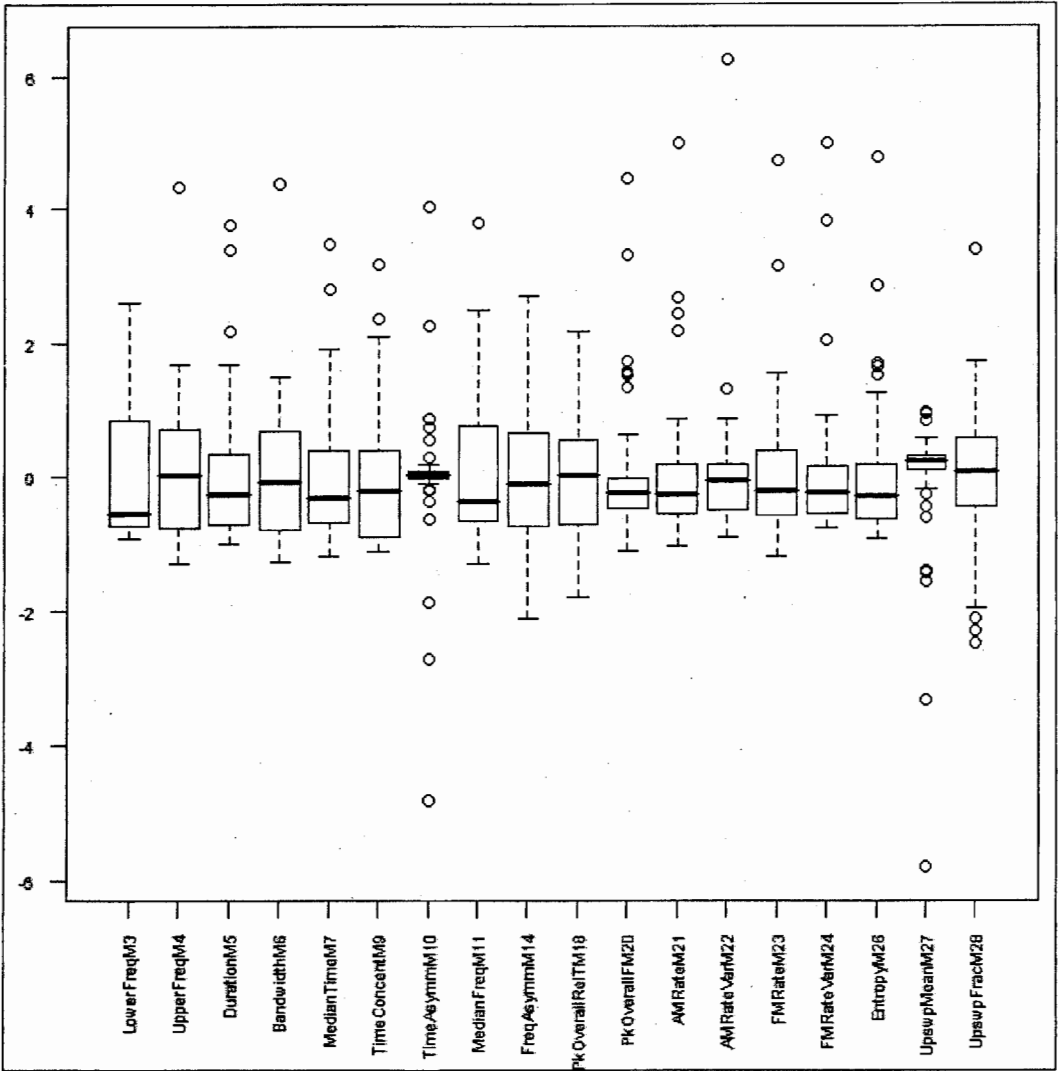


**Figure 53: Boxplot of scaled variables for Bark**  
 Acoustic parameters are on the x-axis. Scaled (normalized) values of the parameters are on the y-axis



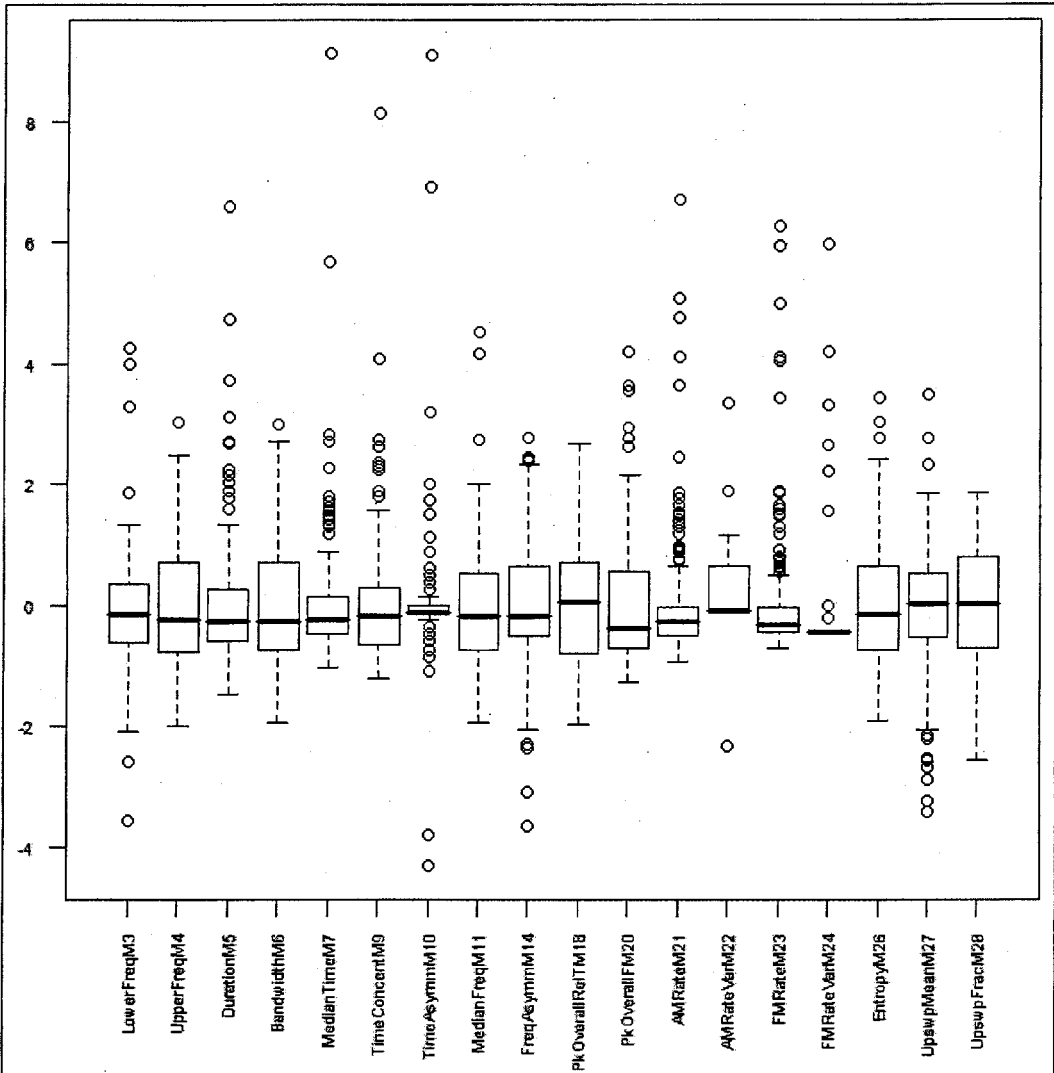
**Figure 54: Boxplot of scaled variables for Roar**

Acoustic parameters are on the x-axis. Scaled (normalized) values of the parameters are on the y-axis



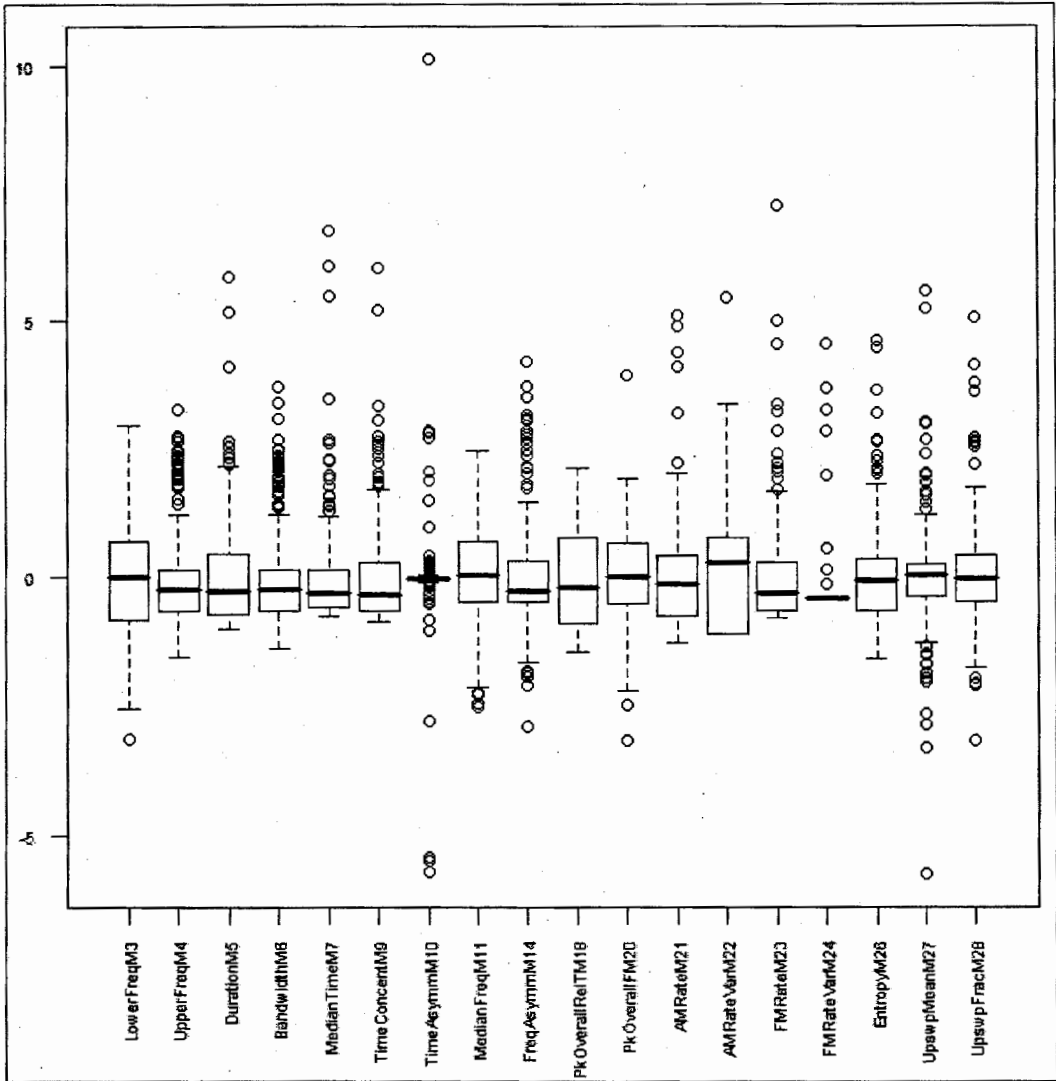
**Figure 55: Boxplot of scaled variables for Rumble**

Acoustic parameters are on the x-axis. Scaled (normalized) values of the parameters are on the y-axis

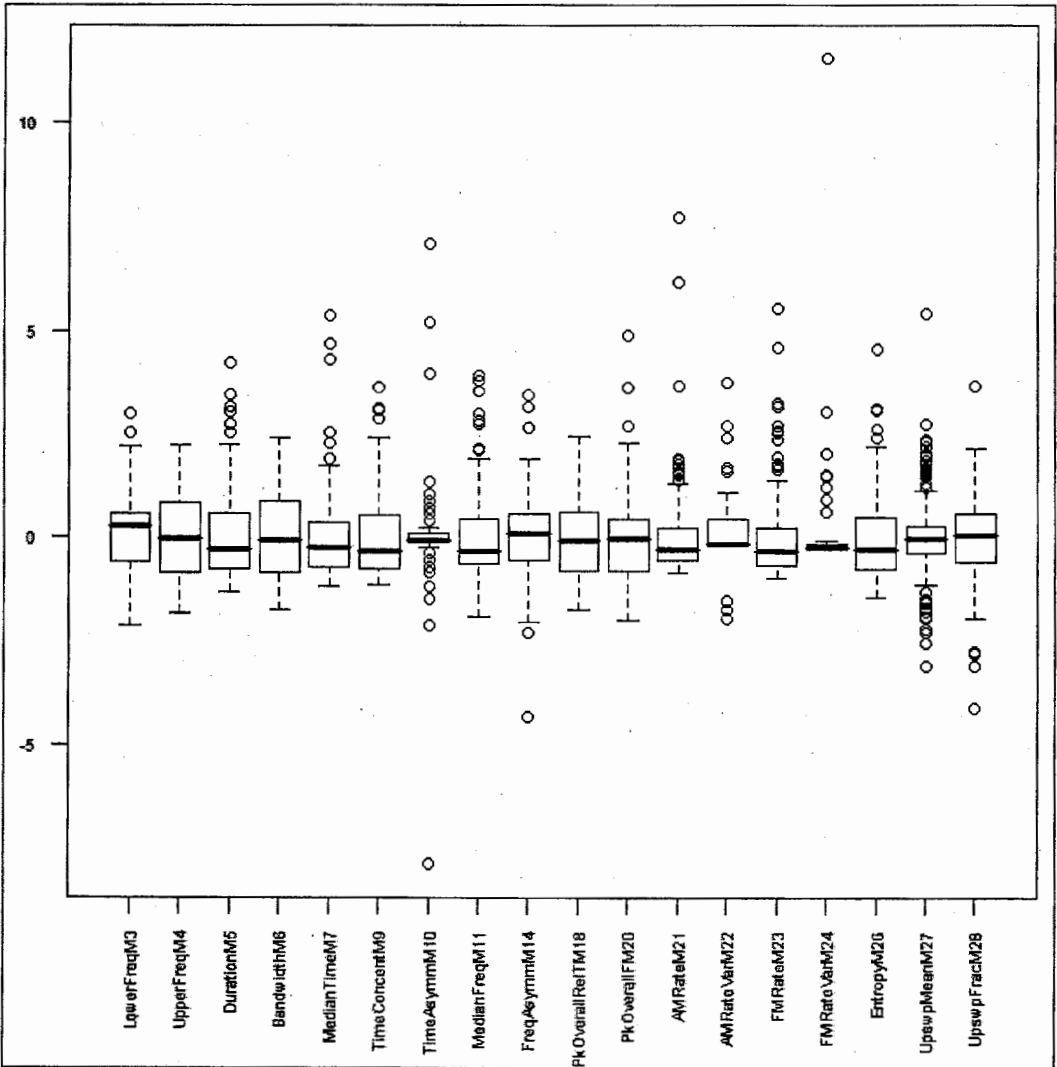


**Figure 56: Boxplot of scaled variables for Squeak**

Acoustic parameters are on the x-axis. Scaled (normalized) values of the parameters are on the y-axis



**Figure 57: Boxplot of scaled variables for Squeal**  
 Acoustic parameters are on the x-axis. Scaled (normalized) values of the parameters are on the y-axis



**Figure 58: Boxplot of scaled variables for Trumpet**

Acoustic parameters are on the x-axis. Scaled (normalized) values of the parameters are on the y-axis



## APPENDIX B: CLASSIFICATION TREE RAW DATA OUTPUT

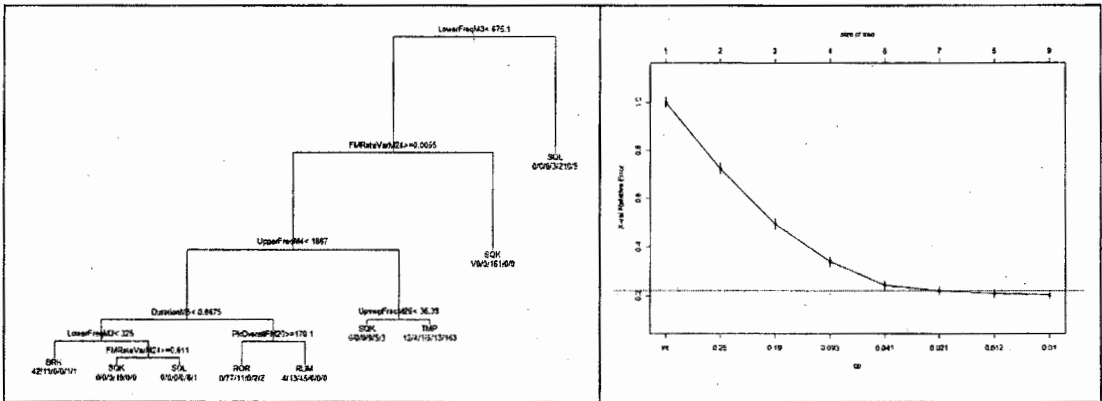


Figure 59: Classification tree final model - plots

Table 12: Classification tree final model - raw data output

```
[1] 0.1220096
> summary(mod)
Call:
rpart(formula = calltype ~ ., data = params)
n = 836
```

	CP	nsplit	rel error	xerror	xstd
1	0.27638191	0	1.0000000	1.0000000	0.02188309
2	0.22948074	1	0.7236181	0.7236181	0.02420221
3	0.16247906	2	0.4941374	0.4941374	0.02314367
4	0.05276382	3	0.3316583	0.3400335	0.02076694
5	0.03182580	5	0.2261307	0.2428811	0.01833773
6	0.01340034	6	0.1943049	0.2211055	0.01766021
7	0.01005025	7	0.1809045	0.2093802	0.01727080
8	0.01000000	8	0.1708543	0.2043551	0.01709825

Node number 1: 836 observations, complexity param=0.2763819  
 predicted class=SQL expected loss=0.7141148  
 class counts: 59 105 57 197 239 179  
 probabilities: 0.071 0.126 0.068 0.236 0.286 0.214  
 left son=2 (614 obs) right son=3 (222 obs)  
 Primary splits:  
 LowerFreqM3 < 675.0855 to the left, improve=163.3921, (0 missing)  
 FMRateVarM24 < 0.0055 to the left, improve=149.4871, (0 missing)  
 MedianFreqM11 < 561.782 to the left, improve=100.2804, (0 missing)  
 PkOverallFM20 < 530.3175 to the left, improve= 99.7010, (0 missing)  
 DurationM5 < 0.4395 to the right, improve= 92.1828, (0 missing)  
 Surrogate splits:  
 PkOverallFM20 < 1302.758 to the left, agree=0.868, adj=0.505, (0 split)  
 MedianFreqM11 < 1318.912 to the left, agree=0.859, adj=0.468, (0 split)  
 DurationM5 < 2.4235 to the left, agree=0.787, adj=0.198, (0 split)  
 AMRateVarM22 < 0.007 to the right, agree=0.785, adj=0.189, (0 split)  
 TimeConcentM9 < 1.6775 to the left, agree=0.782, adj=0.180, (0 split)

Node number 2: 614 observations, complexity param=0.2294807  
 predicted class=SQK expected loss=0.6840391  
 class counts: 59 105 57 194 29 170  
 probabilities: 0.096 0.171 0.093 0.316 0.047 0.277  
 left son=4 (452 obs) right son=5 (162 obs)  
 Primary splits:  
 FMRateVarM24 < 0.0055 to the right, improve=128.59100, (0 missing)  
 AMRateVarM22 < 0.0215 to the right, improve=120.66540, (0 missing)  
 PkOverallFM20 < 530.3175 to the left, improve= 97.40044, (0 missing)

MedianFreqM11 < 561.782 to the left, improve= 97.01046, (0 missing)  
 BandwidthM6 < 2616.284 to the left, improve= 94.27497, (0 missing)  
 Surrogate splits:  
 AMRateVarM22 < 0.0215 to the right, agree=0.879, adj=0.543, (0 split)  
 DurationM5 < 0.287 to the right, agree=0.824, adj=0.333, (0 split)  
 UpswpFracM28 < 34.3755 to the right, agree=0.800, adj=0.241, (0 split)  
 AMRateM21 < 4.0295 to the left, agree=0.796, adj=0.228, (0 split)  
 TimeConcentM9 < 0.1265 to the right, agree=0.787, adj=0.191, (0 split)

Node number 3: 222 observations  
 predicted class=SQL expected loss=0.05405405  
 class counts: 0 0 0 3 210 9  
 probabilities: 0.000 0.000 0.000 0.014 0.946 0.041

Node number 4: 452 observations, complexity param=0.1624791  
 predicted class=TMP expected loss=0.6238938  
 class counts: 58 105 57 33 29 170  
 probabilities: 0.128 0.232 0.126 0.073 0.064 0.376  
 left son=8 (237 obs) right son=9 (215 obs)

Primary splits:  
 UpperFreqM4 < 1866.577 to the left, improve=91.40859, (0 missing)  
 MedianFreqM11 < 642.902 to the left, improve=89.47726, (0 missing)  
 BandwidthM6 < 1361.89 to the left, improve=87.85287, (0 missing)  
 PkOverallFM20 < 468.347 to the left, improve=85.20569, (0 missing)  
 FMRateVarM24 < 0.011 to the right, improve=77.18512, (0 missing)  
 Surrogate splits:  
 BandwidthM6 < 1361.89 to the left, agree=0.985, adj=0.967, (0 split)  
 MedianFreqM11 < 680.2285 to the left, agree=0.889, adj=0.767, (0 split)  
 EntropyM26 < 95.3805 to the left, agree=0.863, adj=0.712, (0 split)  
 PkOverallFM20 < 468.347 to the left, agree=0.858, adj=0.702, (0 split)  
 LowerFreqM3 < 241.5455 to the left, agree=0.827, adj=0.637, (0 split)

Node number 5: 162 observations  
 predicted class=SQK expected loss=0.00617284  
 class counts: 1 0 0 161 0 0  
 probabilities: 0.006 0.000 0.000 0.994 0.000 0.000

Node number 8: 237 observations, complexity param=0.05276382  
 predicted class=ROR expected loss=0.5738397  
 class counts: 46 101 56 19 11 4  
 probabilities: 0.194 0.426 0.236 0.080 0.046 0.017  
 left son=16 (83 obs) right son=17 (154 obs)

Primary splits:  
 DurationM5 < 0.8675 to the left, improve=33.89742, (0 missing)  
 MedianTimeM7 < 0.37 to the right, improve=28.32799, (0 missing)  
 PkOverallFM20 < 170.0515 to the left, improve=23.31349, (0 missing)  
 LowerFreqM3 < 384.906 to the left, improve=22.73504, (0 missing)  
 AMRateVarM22 < 0.3255 to the left, improve=22.52987, (0 missing)  
 Surrogate splits:  
 MedianTimeM7 < 0.4005 to the left, agree=0.920, adj=0.771, (0 split)  
 AMRateM21 < 1.1595 to the right, agree=0.869, adj=0.627, (0 split)  
 FMRateM23 < 1.1525 to the right, agree=0.865, adj=0.614, (0 split)  
 TimeConcentM9 < 0.2685 to the left, agree=0.857, adj=0.590, (0 split)  
 UpperFreqM4 < 1184.326 to the right, agree=0.785, adj=0.386, (0 split)

Node number 9: 215 observations, complexity param=0.01005025  
 predicted class=TMP expected loss=0.227907  
 class counts: 12 4 1 14 18 166  
 probabilities: 0.056 0.019 0.005 0.065 0.084 0.772  
 left son=18 (17 obs) right son=19 (198 obs)

Primary splits:  
 UpswpFracM28 < 36.3945 to the left, improve=11.40962, (0 missing)  
 FMRateVarM24 < 0.011 to the right, improve=11.36771, (0 missing)  
 UpperFreqM4 < 3016.089 to the left, improve=11.33500, (0 missing)  
 TimeConcentM9 < 1.318 to the right, improve=10.79894, (0 missing)  
 BandwidthM6 < 2616.284 to the left, improve=10.19102, (0 missing)  
 Surrogate splits:  
 TimeConcentM9 < 0.0305 to the left, agree=0.930, adj=0.118, (0 split)

TimeAsymmM10 < -0.2115 to the left, agree=0.930, adj=0.118, (0 split)  
DurationM5 < 0.1325 to the left, agree=0.926, adj=0.059, (0 split)  
FMRateM23 < 15.4305 to the right, agree=0.926, adj=0.059, (0 split)

Node number 16: 83 observations, complexity param=0.0318258  
predicted class=BRK expected loss=0.4939759

class counts: 42 11 0 19 9 2  
probabilities: 0.506 0.133 0.000 0.229 0.108 0.024  
left son=32 (55 obs) right son=33 (28 obs)

Primary splits:

LowerFreqM3 < 325.017 to the left, improve=21.43904, (0 missing)  
PkOverallFM20 < 539.676 to the left, improve=18.54354, (0 missing)  
AMRateVarM22 < 0.0215 to the right, improve=18.52093, (0 missing)  
MedianFreqM11 < 669.679 to the left, improve=18.18897, (0 missing)  
DurationM5 < 0.3615 to the right, improve=13.89051, (0 missing)

Surrogate splits:

MedianFreqM11 < 662.2645 to the left, agree=0.940, adj=0.821, (0 split)  
PkOverallFM20 < 539.676 to the left, agree=0.940, adj=0.821, (0 split)  
DurationM5 < 0.3615 to the right, agree=0.904, adj=0.714, (0 split)  
UpperFreqM4 < 1372.742 to the left, agree=0.867, adj=0.607, (0 split)  
AMRateM21 < 2.7815 to the left, agree=0.867, adj=0.607, (0 split)

Node number 17: 154 observations, complexity param=0.05276382  
predicted class=ROR expected loss=0.4155844

class counts: 4 90 56 0 2 2  
probabilities: 0.026 0.584 0.364 0.000 0.013 0.013  
left son=34 (92 obs) right son=35 (62 obs)

Primary splits:

PkOverallFM20 < 170.0515 to the right, improve=28.37610, (0 missing)  
AMRateVarM22 < 0.3255 to the left, improve=24.03463, (0 missing)  
FMRateVarM24 < 0.0695 to the left, improve=22.60592, (0 missing)  
LowerFreqM3 < 22.5425 to the right, improve=18.17635, (0 missing)  
EntropyM26 < 25.117 to the right, improve=15.06470, (0 missing)

Surrogate splits:

MedianFreqM11 < 181.817 to the right, agree=0.857, adj=0.645, (0 split)  
LowerFreqM3 < 35.6645 to the right, agree=0.786, adj=0.468, (0 split)  
AMRateVarM22 < 0.087 to the left, agree=0.753, adj=0.387, (0 split)  
FMRateVarM24 < 0.0195 to the left, agree=0.740, adj=0.355, (0 split)  
UpperFreqM4 < 425.1125 to the right, agree=0.714, adj=0.290, (0 split)

Node number 18: 17 observations

predicted class=SQK expected loss=0.4705882

class counts: 0 0 0 9 5 3  
probabilities: 0.000 0.000 0.000 0.529 0.294 0.176

Node number 19: 198 observations

predicted class=TMP expected loss=0.1767677

class counts: 12 4 1 5 13 163  
probabilities: 0.061 0.020 0.005 0.025 0.066 0.823

Node number 32: 55 observations

predicted class=BRK expected loss=0.2363636

class counts: 42 11 0 0 1 1  
probabilities: 0.764 0.200 0.000 0.000 0.018 0.018

Node number 33: 28 observations, complexity param=0.01340034

predicted class=SQK expected loss=0.3214286

class counts: 0 0 0 19 8 1  
probabilities: 0.000 0.000 0.000 0.679 0.286 0.036  
left son=66 (19 obs) right son=67 (9 obs)

Primary splits:

FMRateVarM24 < 0.011 to the right, improve=11.007940, (0 missing)  
AMRateVarM22 < 0.0215 to the left, improve= 9.385714, (0 missing)  
EntropyM26 < 97.3315 to the right, improve= 8.058442, (0 missing)  
BandwidthM6 < 1012.061 to the right, improve= 4.919048, (0 missing)  
PkOverallRelTM18 < 56.2775 to the left, improve= 4.500000, (0 missing)

Surrogate splits:

AMRateVarM22 < 0.0215 to the left, agree=0.964, adj=0.889, (0 split)

```

EntropyM26 < 80.201 to the right, agree=0.929, adj=0.778, (0 split)
PkOverallRelTM18 < 56.2775 to the left, agree=0.857, adj=0.556, (0 split)
BandwidthM6 < 973.779 to the right, agree=0.821, adj=0.444, (0 split)
MedianFreqM11 < 693.84 to the right, agree=0.821, adj=0.444, (0 split)

```

```

Node number 34: 92 observations
predicted class=ROR expected loss=0.1630435
class counts: 0 77 11 0 2 2
probabilities: 0.000 0.837 0.120 0.000 0.022 0.022

```

```

Node number 35: 62 observations
predicted class=RUM expected loss=0.2741935
class counts: 4 13 45 0 0 0
probabilities: 0.065 0.210 0.726 0.000 0.000 0.000

```

```

Node number 66: 19 observations
predicted class=SQK expected loss=0
class counts: 0 0 0 19 0 0
probabilities: 0.000 0.000 0.000 1.000 0.000 0.000

```

```

Node number 67: 9 observations
predicted class=SQL expected loss=0.1111111
class counts: 0 0 0 0 8 1
probabilities: 0.000 0.000 0.000 0.000 0.889 0.111

```

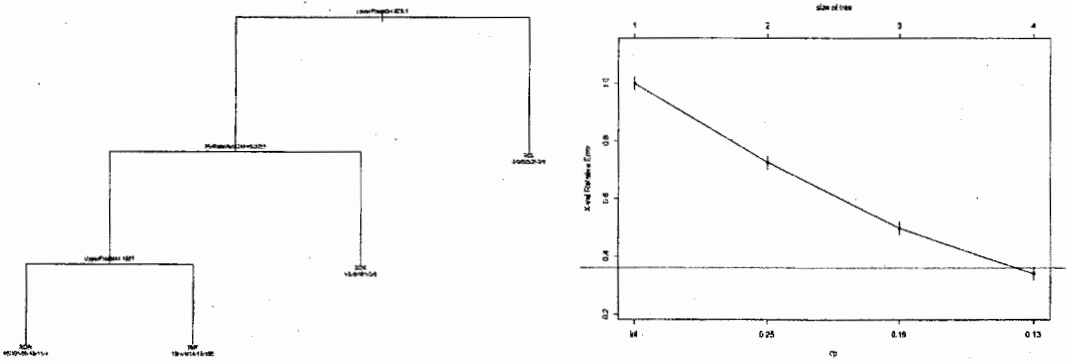


Figure 60: Classification tree validation model - plots

Table 13: Classification tree validation model - raw data output

```

[1] 0.2368421
> summary(new)
Call:
rpart(formula = calltype ~ ., data = params)
n = 836

      CP nsplit rel error   xerror   xstd
1 0.2763819      0 1.0000000 1.0000000 0.02188309
2 0.2294807      1 0.7236181 0.7252931 0.02420020
3 0.1624791      2 0.4941374 0.4974874 0.02317902
4 0.1000000      3 0.3316583 0.3400335 0.02076694

Node number 1: 836 observations, complexity param=0.2763819
predicted class=SQL expected loss=0.7141148
class counts: 59 105 57 197 239 179
probabilities: 0.071 0.126 0.068 0.236 0.286 0.214
left son=2 (614 obs) right son=3 (222 obs)
Primary splits:
LowerFreqM3 < 675.0855 to the left, improve=163.3921, (0 missing)
FMRateVarM24 < 0.0055 to the left, improve=149.4871, (0 missing)
MedianFreqM11 < 561.782 to the left, improve=100.2804, (0 missing)

```

PkOverallFM20 < 530.3175 to the left, improve= 99.7010, (0 missing)  
 DurationM5 < 0.4395 to the right, improve= 92.1828, (0 missing)  
 Surrogate splits:  
 PkOverallFM20 < 1302.758 to the left, agree=0.868, adj=0.505, (0 split)  
 MedianFreqM11 < 1318.912 to the left, agree=0.859, adj=0.468, (0 split)  
 DurationM5 < 2.4235 to the left, agree=0.787, adj=0.198, (0 split)  
 AMRateVarM22 < 0.007 to the right, agree=0.785, adj=0.189, (0 split)  
 TimeConcentM9 < 1.6775 to the left, agree=0.782, adj=0.180, (0 split)

Node number 2: 614 observations, complexity param=0.2294807  
 predicted class=SQK expected loss=0.6840391  
 class counts: 59 105 57 194 29 170  
 probabilities: 0.096 0.171 0.093 0.316 0.047 0.277  
 left son=4 (452 obs) right son=5 (162 obs)

Primary splits:  
 FMRateVarM24 < 0.0055 to the right, improve=128.59100, (0 missing)  
 AMRateVarM22 < 0.0215 to the right, improve=120.66540, (0 missing)  
 PkOverallFM20 < 530.3175 to the left, improve= 97.40044, (0 missing)  
 MedianFreqM11 < 561.782 to the left, improve= 97.01046, (0 missing)  
 BandwidthM6 < 2616.284 to the left, improve= 94.27497, (0 missing)  
 Surrogate splits:  
 AMRateVarM22 < 0.0215 to the right, agree=0.879, adj=0.543, (0 split)  
 DurationM5 < 0.287 to the right, agree=0.824, adj=0.333, (0 split)  
 UpswpFracM28 < 34.3755 to the right, agree=0.800, adj=0.241, (0 split)  
 AMRateM21 < 4.0295 to the left, agree=0.796, adj=0.228, (0 split)  
 TimeConcentM9 < 0.1265 to the right, agree=0.787, adj=0.191, (0 split)

Node number 3: 222 observations  
 predicted class=SQL expected loss=0.05405405  
 class counts: 0 0 0 3 210 9  
 probabilities: 0.000 0.000 0.000 0.014 0.946 0.041

Node number 4: 452 observations, complexity param=0.1624791  
 predicted class=TMP expected loss=0.6238938  
 class counts: 58 105 57 33 29 170  
 probabilities: 0.128 0.232 0.126 0.073 0.064 0.376  
 left son=8 (237 obs) right son=9 (215 obs)

Primary splits:  
 UpperFreqM4 < 1866.577 to the left, improve=91.40859, (0 missing)  
 MedianFreqM11 < 642.902 to the left, improve=89.47726, (0 missing)  
 BandwidthM6 < 1361.89 to the left, improve=87.85287, (0 missing)  
 PkOverallFM20 < 468.347 to the left, improve=85.20569, (0 missing)  
 FMRateVarM24 < 0.011 to the right, improve=77.18512, (0 missing)  
 Surrogate splits:  
 BandwidthM6 < 1361.89 to the left, agree=0.985, adj=0.967, (0 split)  
 MedianFreqM11 < 680.2285 to the left, agree=0.889, adj=0.767, (0 split)  
 EntropyM26 < 95.3805 to the left, agree=0.863, adj=0.712, (0 split)  
 PkOverallFM20 < 468.347 to the left, agree=0.858, adj=0.702, (0 split)  
 LowerFreqM3 < 241.5455 to the left, agree=0.827, adj=0.637, (0 split)

Node number 5: 162 observations  
 predicted class=SQK expected loss=0.00617284  
 class counts: 1 0 0 161 0 0  
 probabilities: 0.006 0.000 0.000 0.994 0.000 0.000

Node number 8: 237 observations  
 predicted class=ROR expected loss=0.5738397  
 class counts: 46 101 56 19 11 4  
 probabilities: 0.194 0.426 0.236 0.080 0.046 0.017

Node number 9: 215 observations  
 predicted class=TMP expected loss=0.227907  
 class counts: 12 4 1 14 18 166  
 probabilities: 0.056 0.019 0.005 0.065 0.084 0.772

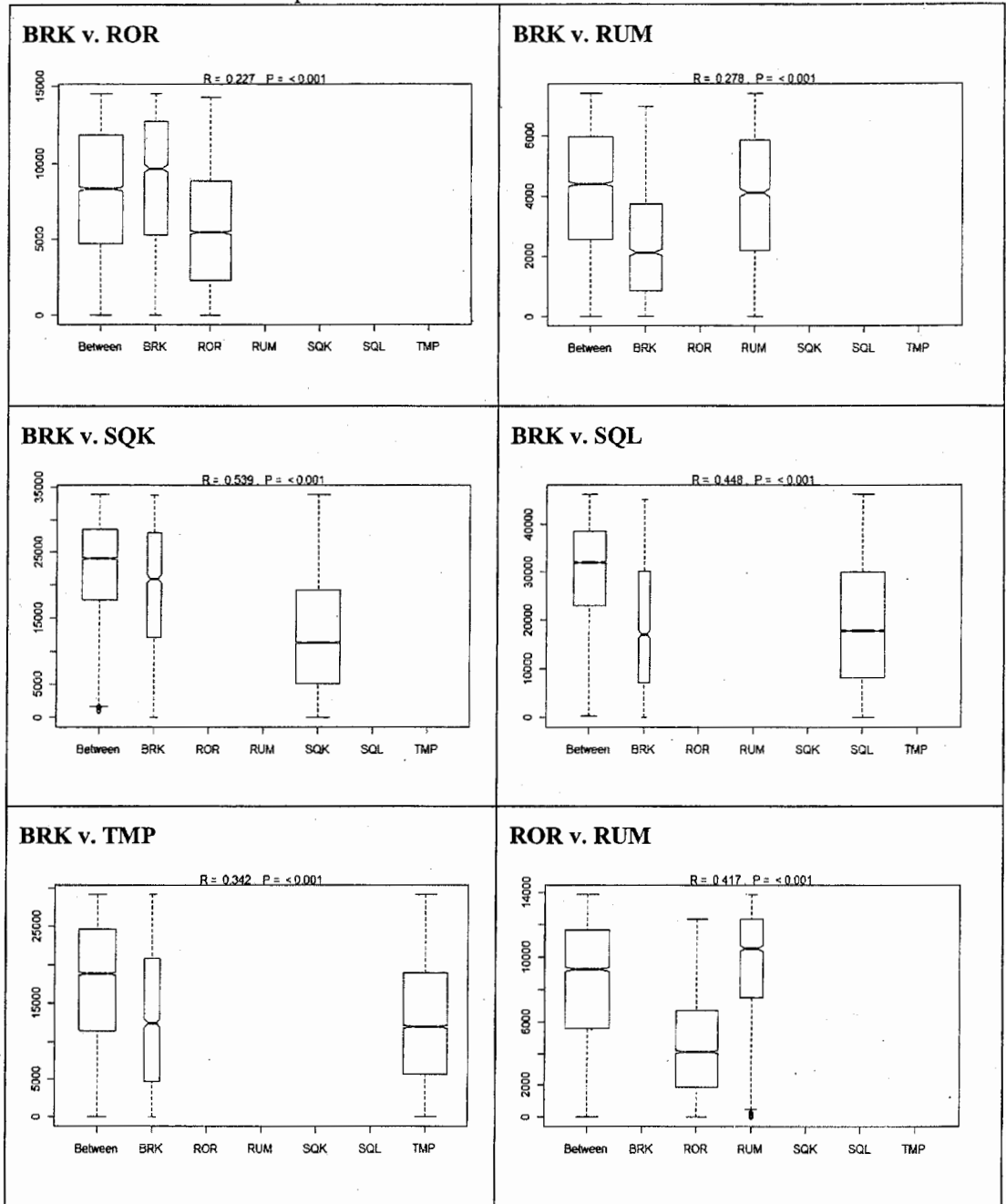
## APPENDIX C: PRINCIPAL COMPONENT ANALYSIS DATA – RAW DATA OUTPUT

**Table 14: Principal Component Analysis - raw data output**

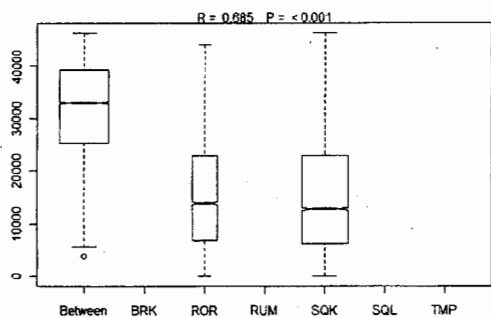
Standard deviation	2.1032890	1.8106105	1.31698516	1.15100919	1.11924231
Proportion of Variance	0.2460624	0.1823465	0.09647373	0.07368938	0.06967798
Cumulative Proportion	0.2460624	0.4284088	0.52488258	0.59857195	0.66824993
 Loadings:					
	Comp.1	Comp.2	Comp.3	Comp.4	Comp.5
LowerFreqM3	-0.29	-0.14	-0.35	-0.17	0.14
UpperFreqM4	-0.42		0.25	0.15	
DurationM5		-0.52			
BandwidthM6	-0.38		0.35	0.20	
MedianTimeM7		-0.51		0.18	
TimeConcentM9		-0.51	-0.11	0.12	
TimeAsymmM10				-0.32	0.35
MedianFreqM11	-0.44				0.15
FreqAsymmM14			0.33	0.20	-0.23
PkOverallRelTM18				0.28	-0.47
PkOverallFM20	-0.38	-0.10	-0.24	-0.13	0.14
AMRateM21		0.24	-0.34	0.17	
AMRateVarM22	0.23		0.28	0.12	0.42
FMRateM23		0.24	-0.25	0.10	
FMRateVarM24	0.21		0.24		0.46
EntropyM26	-0.38		0.24	0.15	
UpswpMeanM27			0.17	-0.54	-0.29
UpswpFracM28		-0.16	0.31	-0.48	-0.25
 broken.stick(18)					
	j	E(j)			
[1,]	1	0.194172671			
[2,]	2	0.138617115			
[3,]	3	0.110839338			
[4,]	4	0.092320819			
[5,]	5	0.078431930			
[6,]	6	0.067320819			
[7,]	7	0.058061560			
[8,]	8	0.050125052			
[9,]	9	0.043180608			
[10,]	10	0.037007768			
[11,]	11	0.031452212			
[12,]	12	0.026401707			
[13,]	13	0.021772078			
[14,]	14	0.017498574			
[15,]	15	0.013530320			
[16,]	16	0.009826616			
[17,]	17	0.006354394			
[18,]	18	0.003086420			
		Comp.1	Comp.2	Comp.3	Comp.4
Comp.6					Comp.5
Proportion of Variance	0.05750879	0.2460624	0.1823465	0.09647373	0.07368938
		0.06967798			

# APPENDIX D: ANALYSIS OF SIMILARITY (ANOSIM) DATA OUTPUT

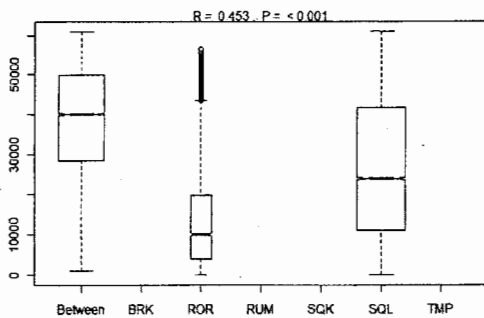
**Table 15: Pair-wise ANOSIM plots for 15 pair-wise tests**  
 Pairs are noted above each plot.



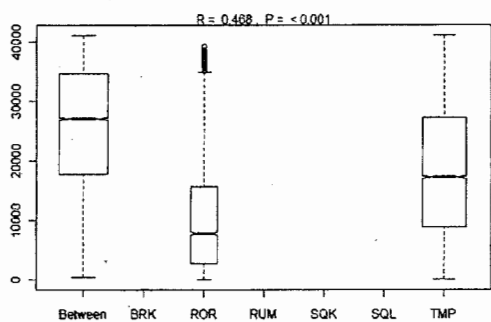
### ROR v. SQK



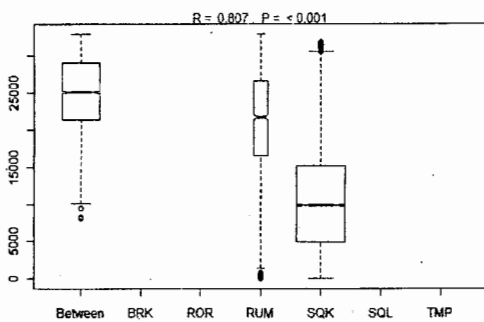
### ROR v. SQL



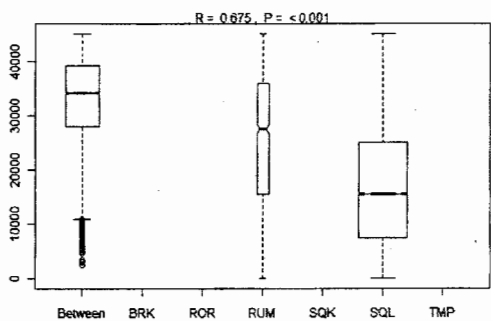
### ROR v. TMP



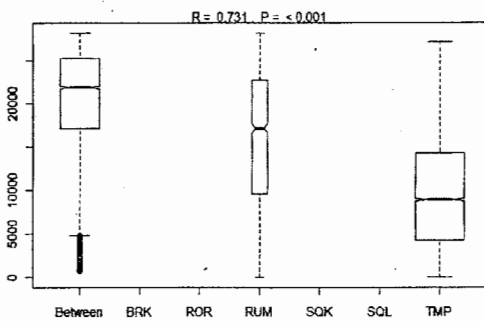
### RUM v. SQK



### RUM v. SQL

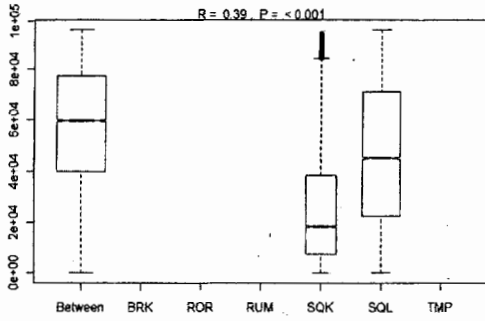


### RUM v. TMP

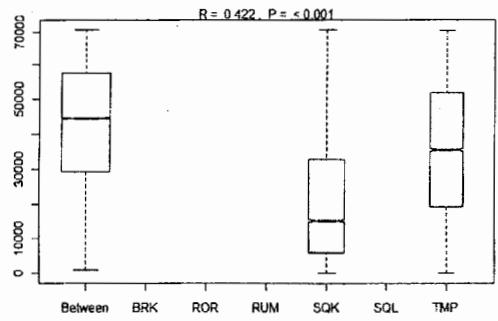




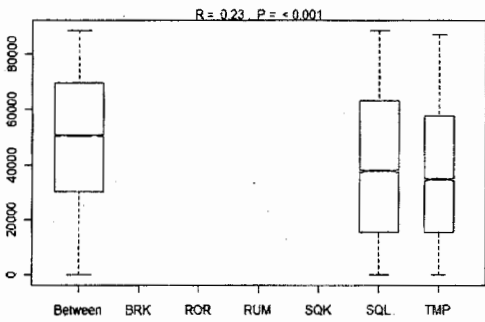
### SQK v. SQL



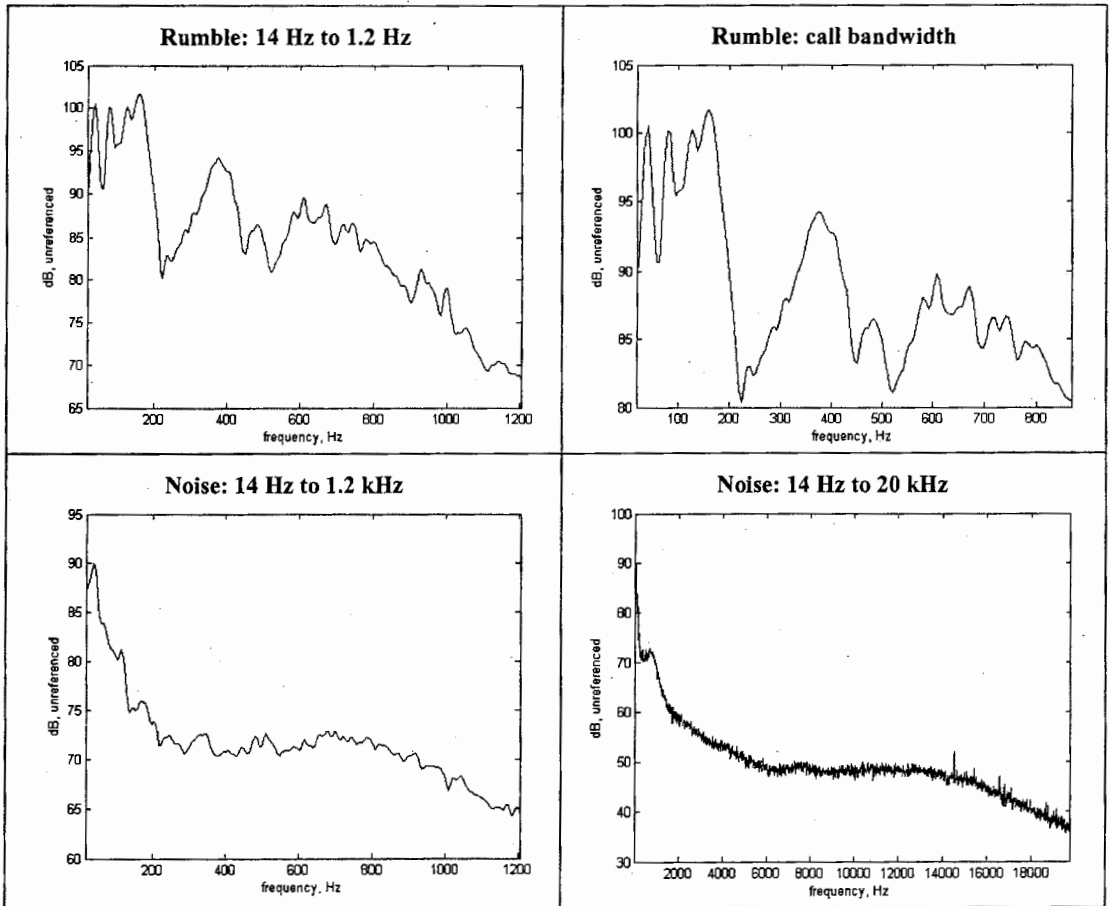
### SQK v. TMP



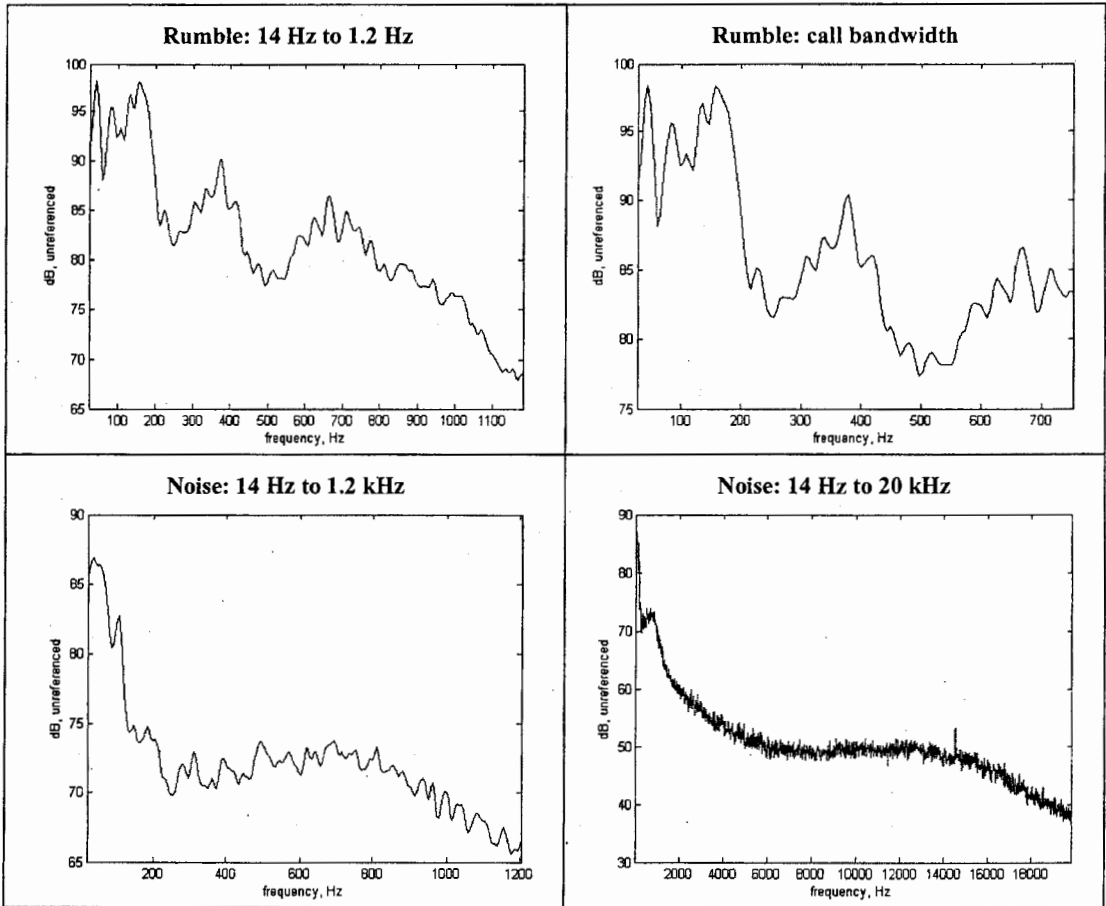
### SQL v. TMP



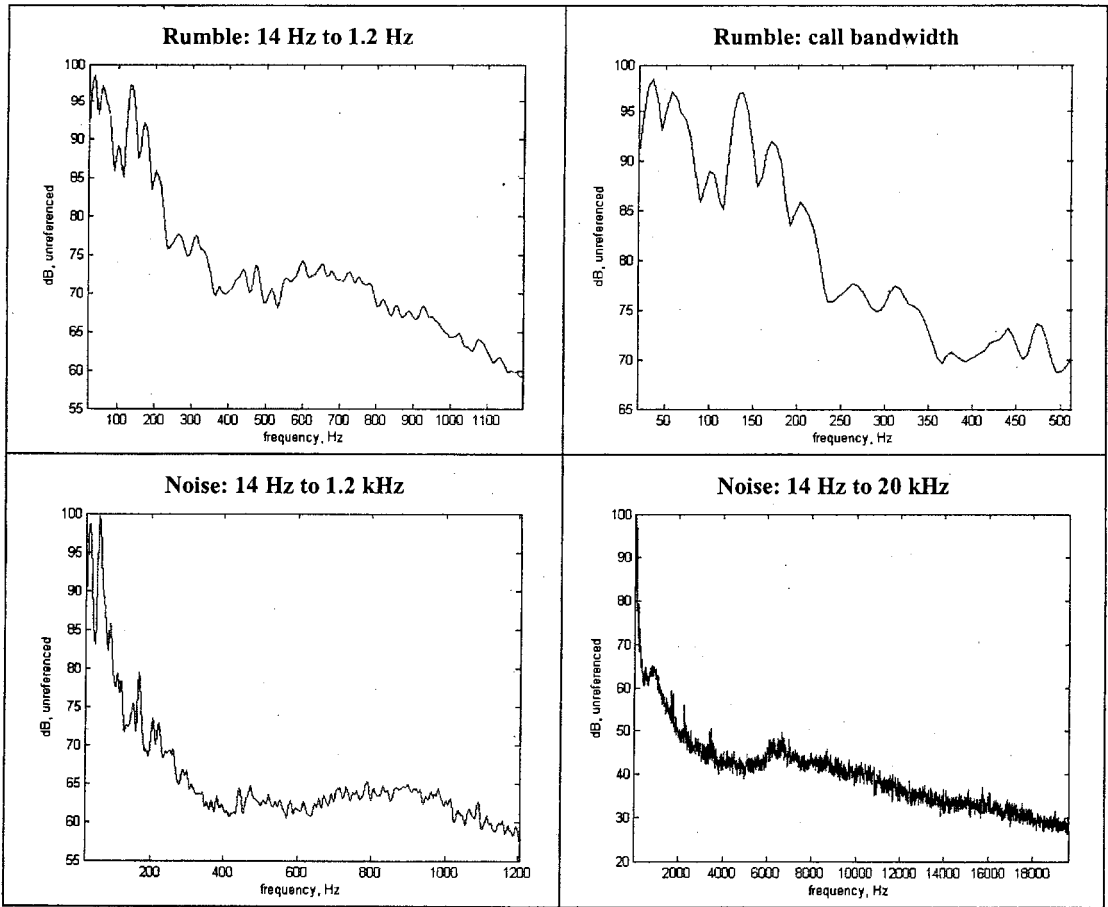
## APPENDIX E: POWER SPECTRA OF RUMBLES AND NOISE



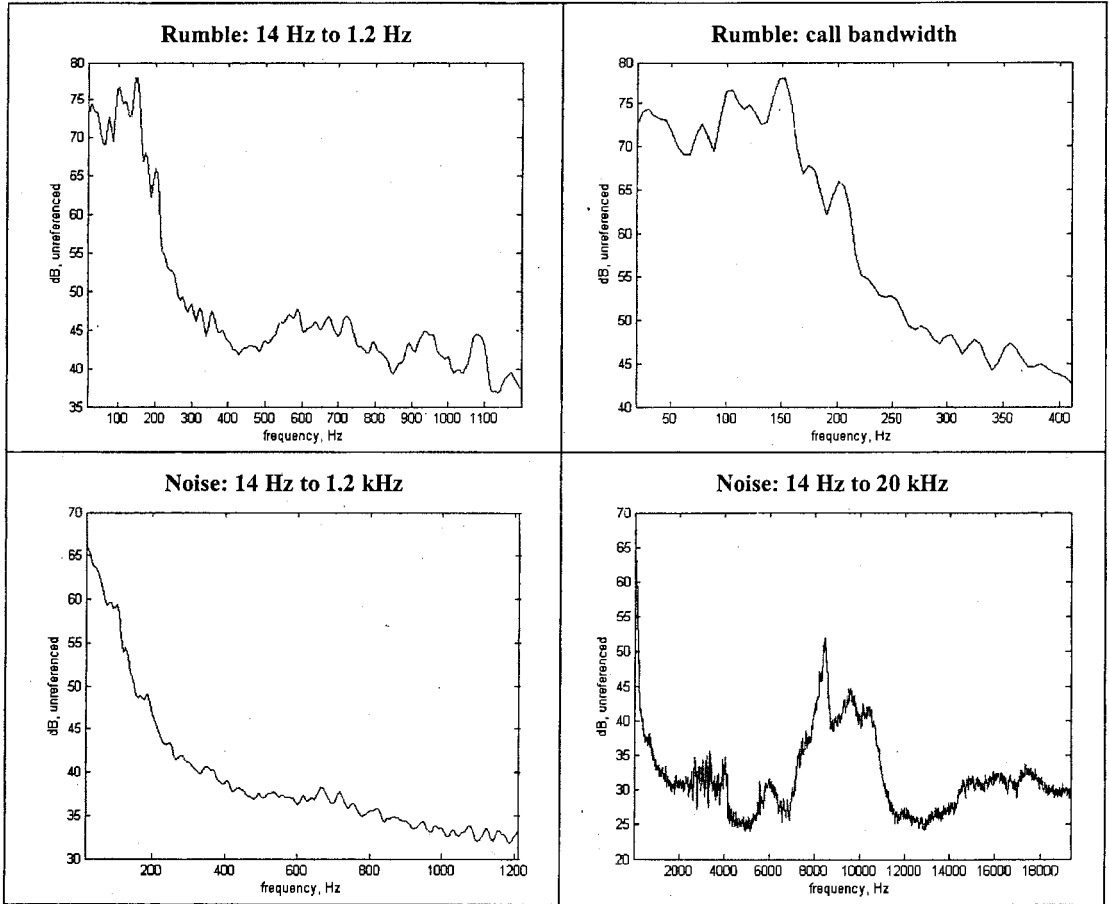
**Figure 61: Power spectra of Rumble and background noise at the Oregon Zoo (example 1)**  
Frequency is on the x-axis. Power (dB) is on the y-axis. Power is referenced to the same value.



**Figure 62: Power spectra of Rumble and background noise at the Oregon Zoo (example 2)**  
 Frequency is on the x-axis. Power (dB) is on the y-axis. Power is referenced to the same value.

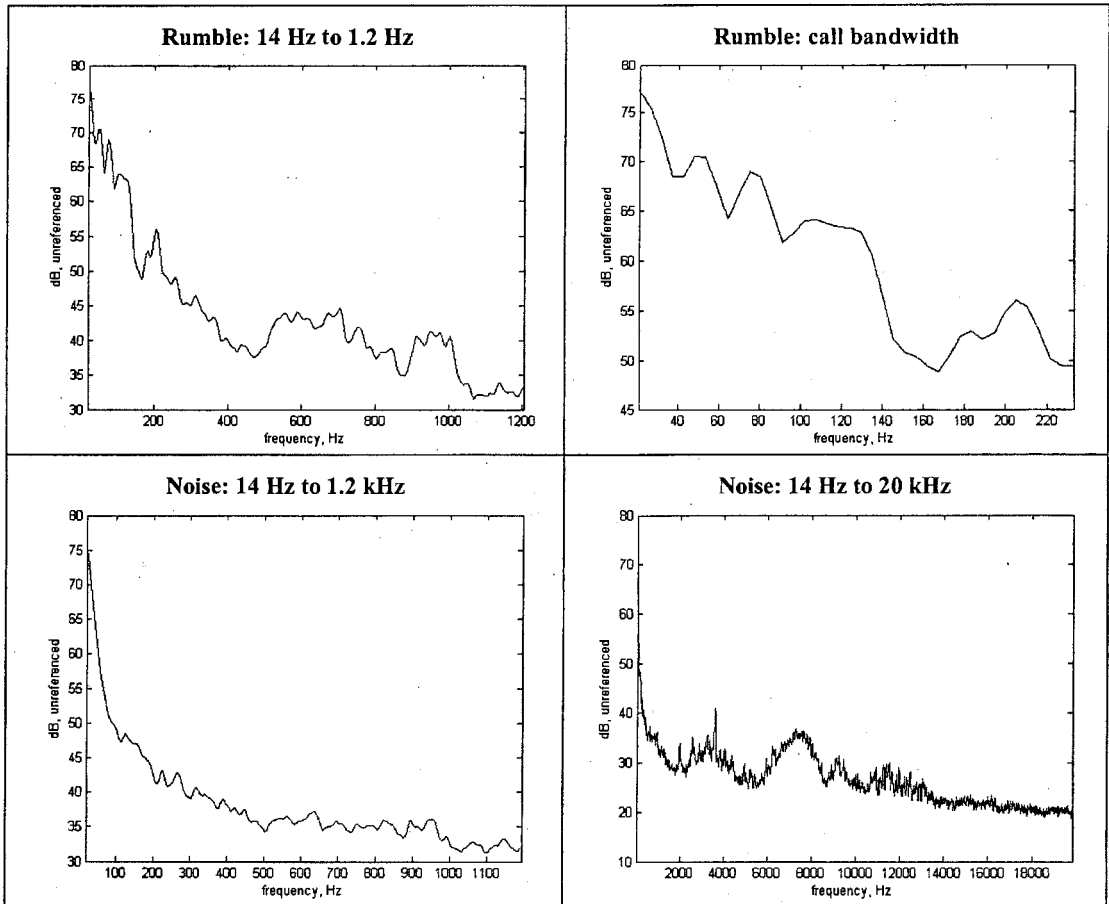


**Figure 63: Power spectra of Rumble and background noise at the Oregon Zoo (example 3)**  
 Frequency is on the x-axis. Power (dB) is on the y-axis. Power is referenced to the same value.

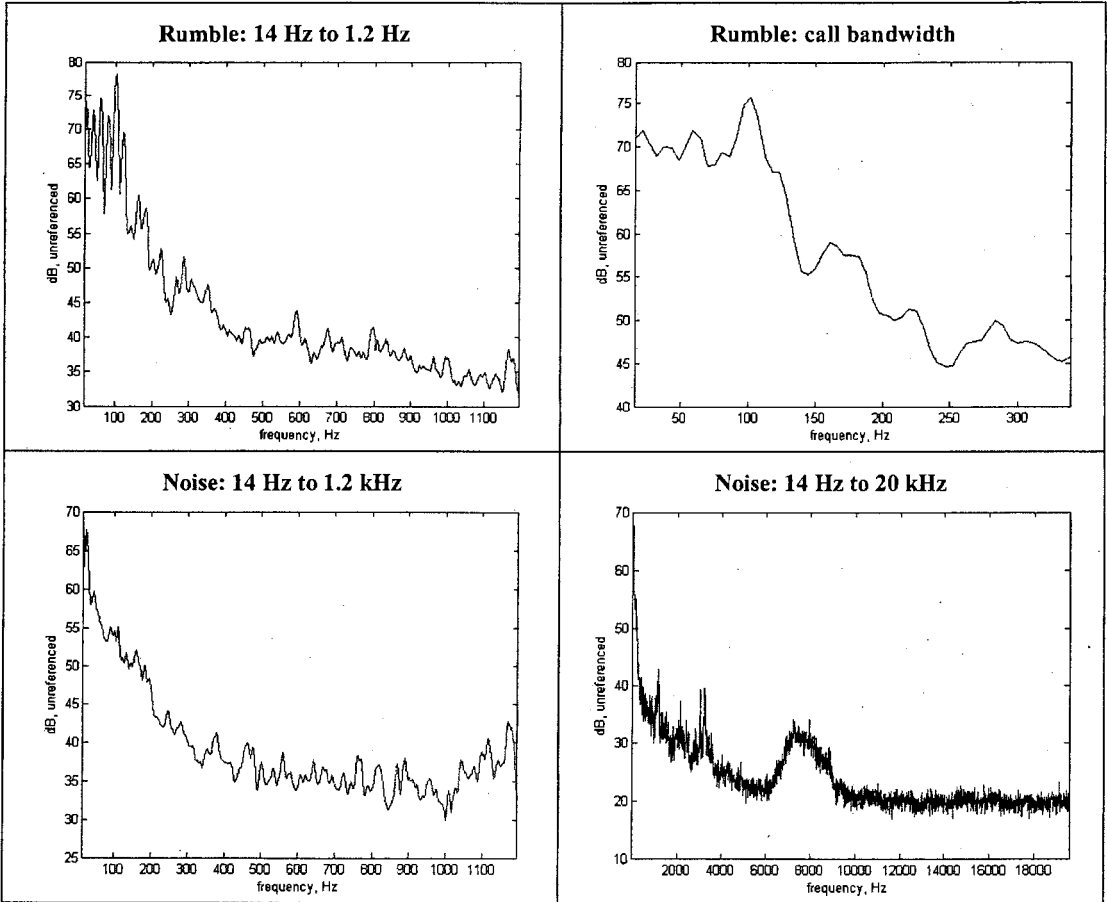


**Figure 64: Power spectra of Rumble and background noise at the Elephant Nature Park (example 1)**

Frequency is on the x-axis. Power (dB) is on the y-axis. Power is referenced to the same value.

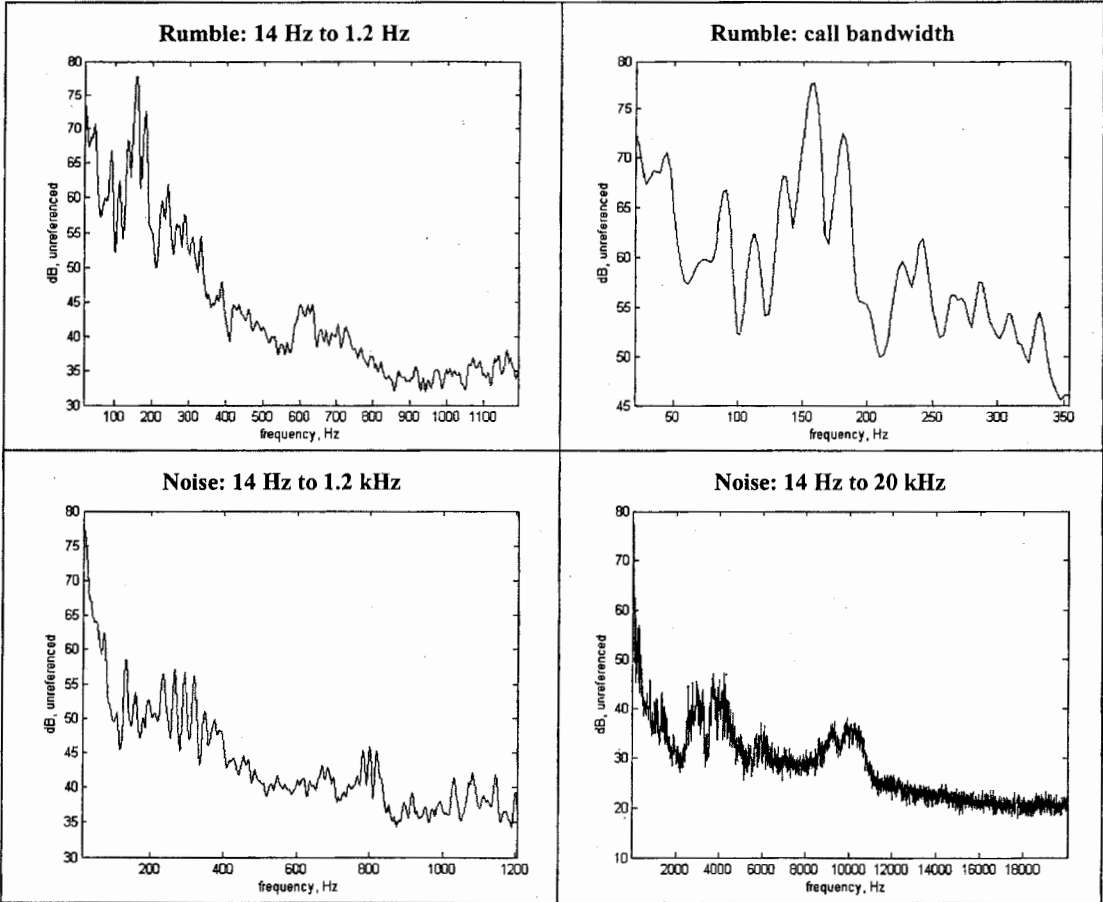


**Figure 65: Power spectra of Rumble and background noise at the Elephant Nature Park (example 2)**  
 Frequency is on the x-axis. Power (dB) is on the y-axis. Power is referenced to the same value.



**Figure 66: Power spectra of Rumble and background noise at the Elephant Nature Park (example 3)**

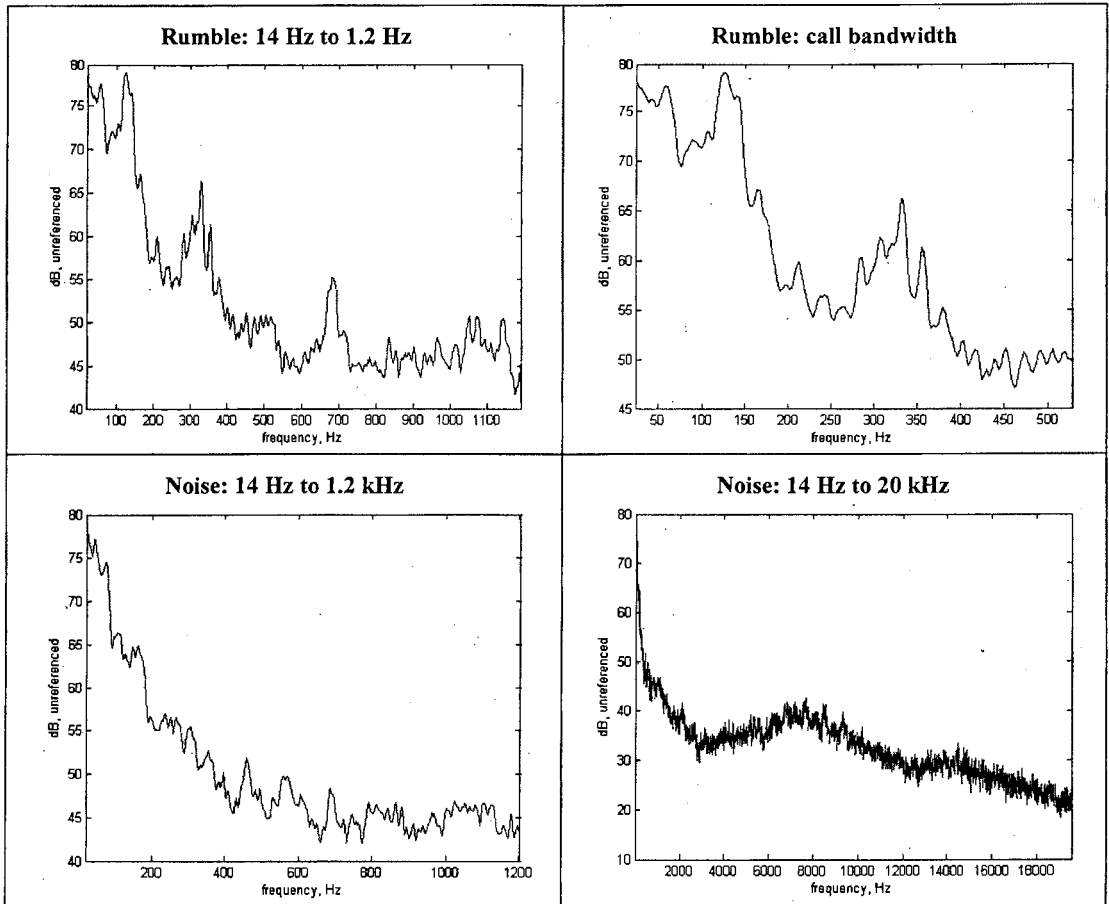
Frequency is on the x-axis. Power (dB) is on the y-axis. Power is referenced to the same value.



**Figure 67: Power spectra of Rumble and background noise at the Elephant Nature Park (example 4)**

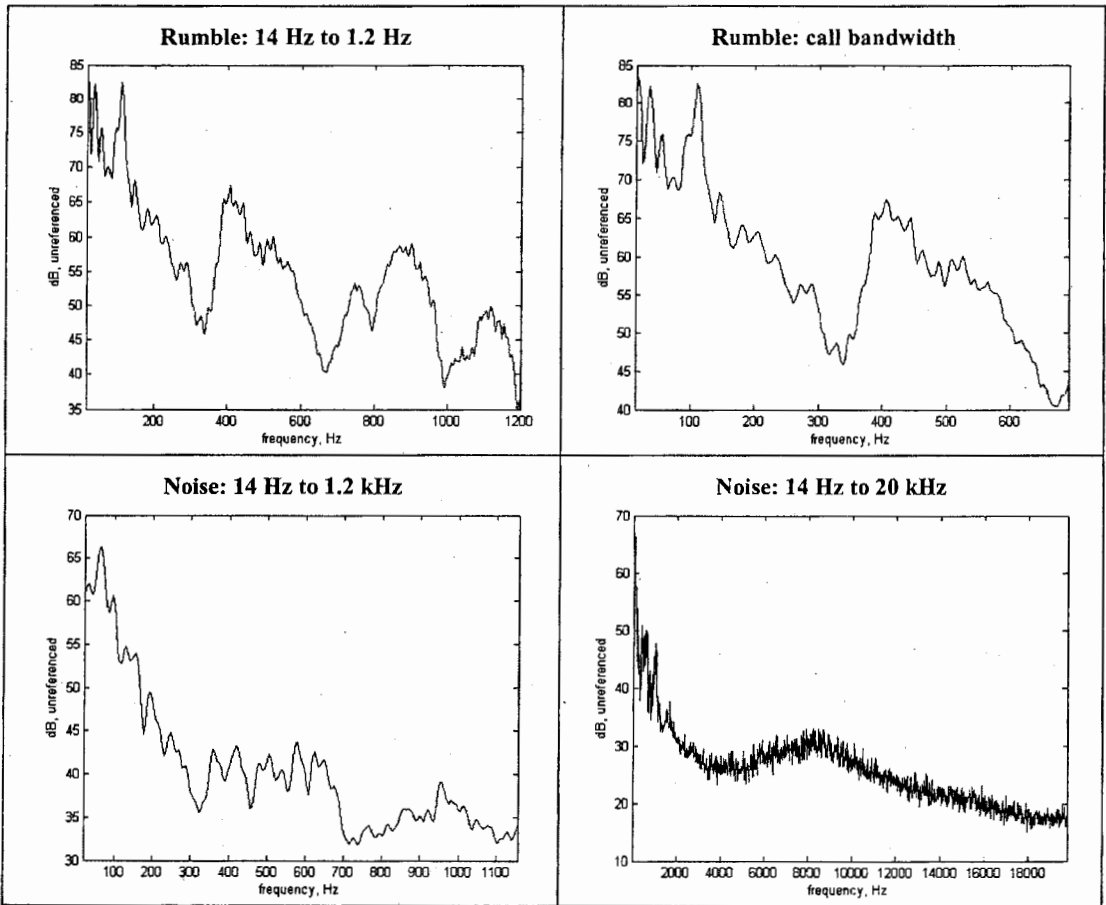
Frequency is on the x-axis. Power (dB) is on the y-axis. Power is referenced to the same value.





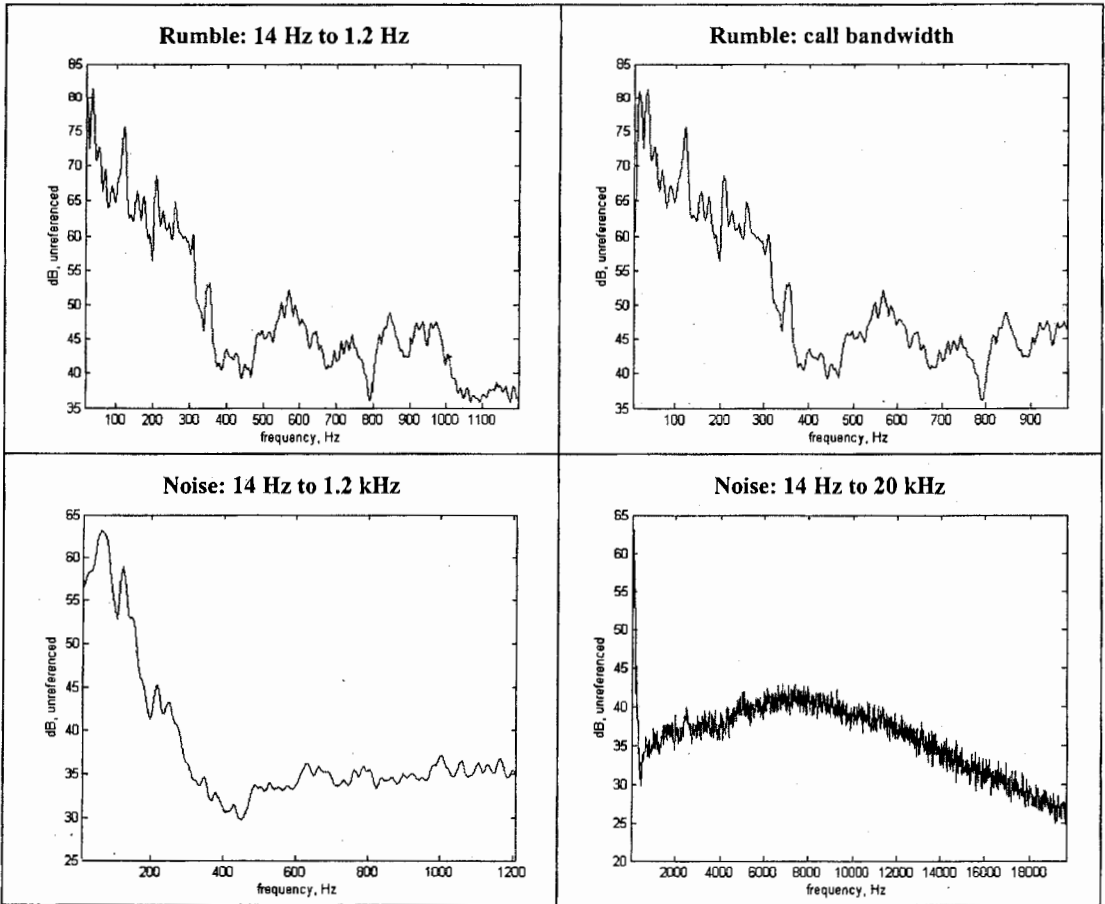
**Figure 68: Power spectra of Rumble and background noise at the Royal Elephant Kraal (example 1)**

Frequency is on the x-axis. Power (dB) is on the y-axis. Power is referenced to the same value.



**Figure 69: Power spectra of Rumble and background noise at the Royal Elephant Kraal (example 2)**

Frequency is on the x-axis. Power (dB) is on the y-axis. Power is referenced to the same value.



**Figure 70: Power spectra of Rumble and background noise at the Royal Elephant Kraal (example 3)**

Frequency is on the x-axis. Power (dB) is on the y-axis. Power is referenced to the same value.