

An empirical investigation of loudness fluctuations in South African broadcast audio

by

Jozua Johannes Georg Loots

A mini-dissertation submitted in partial fulfilment of the requirements for the degree
Master of Music (Music Technology)

in the Department of Music at the
UNIVERSITY OF PRETORIA
FACULTY OF HUMANITIES

Supervisor: Mr W.M. de Villiers

October 2015

Acknowledgements

First and foremost, I need to thank my supervisor Murray de Villiers for not only guiding me through this research project, but for being a mentor, and for igniting my passion for audio technology in the first place.

Special thanks to Professor Leon Janikian and Professor Jim Anderson at Northeastern University, without whom I never would have embarked on this degree, and career path in the first place. The management at the Open Window Institute, especially Pluto Panoussis, for the support needed in the final stages of this project. To all the members of the PLOUD research group at the EBU for the amazing research, for achieving impressive results across global, political and audio-ideological boundaries; especially Florian Camerer for his passionate work in loudness advocacy – bringing us a step closer to loudness nirvana. The developers at Steinberg for including an impressive loudness toolset in the latest versions of Nuendo, and for their extremely rapid response to technical queries.

Lastly, I would like to thank Shanna Nienaber for her unwavering support, sound advice, endless enthusiasm, and perfect cups of tea throughout this entire project.

Table of contents

Acknowledgements	i
List of tables	iv
List of figures	v
List of abbreviations and acronyms	vii
Terminology and concept clarification	ix
1. Introduction: loudness fluctuations in broadcast audio	1
1.1 Introduction and identification of the research theme.....	1
1.2 Identification and demarcation of the research question.....	3
1.3 Delimitation, purpose and value of the study.....	4
1.4 Structure of the research.....	5
2. Theoretical considerations: audio measurement and perception	6
2.1 Audio levelling.....	6
2.1.1 Dynamic range, and medium limitations.....	6
2.1.2 The language of audio level descriptors.....	7
2.1.3 Level measurement and metering.....	8
2.1.4 Broadcast audio levelling standards.....	9
2.1.5 Dynamic range manipulation, and the ‘Loudness War’.....	10
2.1.6 Conclusion.....	12
2.2 Measuring loudness.....	13
2.2.1 Psychoacoustics and loudness perception.....	13
2.2.2 International standardisation of loudness measurement.....	15
2.2.3 Universal loudness descriptors and EBU R128.....	18
2.2.4 Conclusion: South African adoption of loudness normalisation.....	19
2.3 Listener tolerance.....	19
2.3.1 Listening ‘comfort zone’.....	20
2.3.2 Dynamic range tolerance.....	21
2.3.3 Conclusion: loudness fluctuation scenarios.....	22
2.4 Conclusion.....	23
3. Experimental methodology	24
3.1 Technical considerations.....	24
3.1.1 Hardware.....	24
3.1.2 Loudness metering and logging.....	24
3.1.3 Calibration.....	25
3.2 Sample selection.....	25
3.2.1 Channel and station selection.....	25

3.2.2 Time and genre sample selection.....	26
3.2.3 Radio Zap Tests Simulations	27
3.3 Ethical and legal considerations	28
3.4 Data analysis.....	28
4. High level research findings	30
4.1 Radio findings	30
4.2 Television findings.....	33
5. Detailed findings, analysis and discussion	35
5.1 Intra-station fluctuations in radio loudness.....	35
5.2 Discussion of individual radio stations.....	37
5.2.1 Ukhozi FM.....	37
5.2.2 Metro FM.....	38
5.2.3 UW FM.....	40
5.2.4 Lesedi FM.....	41
5.2.5 Thobela FM	42
5.2.6 Motsweding	44
5.2.7 Jacaranda	45
5.2.8 RSG	46
5.2.9 5FM.....	48
5.2.10 Ikwekwezi FM	49
5.2.13 Classic FM	51
5.3 Zap Tests	51
5.4 Zap Testing as efficient methodology for determining station loudness.....	55
5.5 Loudness Range values as fluctuation indicators.....	57
5.6 Conclusion	58
6. Conclusion and recommendations	59
6.1 Summary of findings	59
6.2 Recommendations for broadcasters and content producers.....	61
6.3 Suggestions for further research	63
7. List of References	64
Appendix A – Loudness model comparisons	69
Appendix B – Full Radio Audience Measurement Survey (RAMS)	70
Appendix C – Full Programme Loudness Distribution Graphs.....	71
Summary.....	74
Keywords.....	75

List of tables

Table 1: Digital medium dynamic range in terms of full-scale signal to quantisation noise	7
Table 2: The effect of inter-program loudness fluctuations on the audience	21
Table 3: National Radio Audience Measurement Survey for May 2014 - June 2015: Top 10	26
Table 4: Order of FM radio stations according to frequency for ‘zap test’ simulation	28
Table 5: Station integrated loudness and loudness range values for top 10 radio stations.....	30
Table 6: Station loudness and loudness range for TV stations.....	34
Table 7: Percentage of fluctuations outside ‘comfort zone’ – top 10 radio stations	35
Table 8: Average fluctuation magnitude for zap tests – all radio stations.....	53
Table 9: Radio Audience Measurement Survey for May 2014 - June 2015	70

List of figures

Figure 1.1: Peak-level normalisation vs. loudness normalisation of audio.....	2
Figure 2.1: Programme alignment in the analogue world.....	9
Figure 2.2: Broadcast peak programme meter.....	10
Figure 2.3: Frequency spacing of FM channels.....	11
Figure 2.4: Equal-loudness contours according to ISO226:2003.....	13
Figure 2.5: A comparison of objective and subjective loudness for the ITU loudness meter.....	16
Figure 2.6: Channel processing and summation in ITU-R BS.1770.....	17
Figure 2.7: Listener loudness ‘comfort zone’.....	20
Figure 2.8: Dynamic Range Tolerance for various listening situations.....	21
Figure 3.1: Example single “zap test” simulation of the top 10 radio stations.....	27
Figure 4.1: Station integrated loudness with genre breakdown for top 10 radio stations.....	31
Figure 4.2: Station loudness range with genre breakdown for top 10 radio stations.....	32
Figure 4.3: Station loudness distributions for TV stations.....	33
Figure 4.4: Waveform example of automatic gain control in television broadcast.....	34
Figure 5.1: Proportion of fluctuations outside listener ‘comfort zone’.....	36
Figure 5.3: Ukhozi FM Fluctuation Distribution.....	37
Figure 5.4: Ukhozi FM Loudness Range Distribution.....	38
Figure 5.5: Metro FM Fluctuation Distribution.....	39
Figure 5.6: Metro FM Loudness Range Distribution.....	39
Figure 5.7: UW FM Fluctuation Distribution.....	40
Figure 5.8: UW FM Loudness Range Distribution.....	40
Figure 5.9: Lesedi FM Fluctuation Distribution.....	41
Figure 5.10: Lesedi FM Loudness Range Distribution.....	42
Figure 5.11: Thobela FM Fluctuation Distribution.....	43
Figure 5.12: Thobela FM Loudness Range Distribution.....	43
Figure 5.13: Motswedding Fluctuation Distribution.....	44
Figure 5.14: Motswedding Loudness Range Distribution.....	45
Figure 5.15: Jacaranda Fluctuation Distribution.....	45
Figure 5.16: Jacaranda Loudness Range Distribution.....	46
Figure 5.17: RSG Fluctuation Distribution.....	46
Figure 5.18: RSG Loudness Range Distribution.....	47
Figure 5.19: 5FM Fluctuation Distribution.....	48
Figure 5.20: 5FM Loudness Range Distribution.....	49
Figure 5.21: Ikwekwezi FM Fluctuation Distribution.....	49
Figure 5.22: Ikwekwezi FM Loudness Range Distribution.....	50

Figure 5.23: Classic FM loudness distribution of adjacent programme segments	51
Figure 5.24: Integrated loudness from zap test – all stations	52
Figure 5.25: Proportion of fluctuations outside listener ‘comfort zone’ – zap tests.....	54
Figure 5.26: Integrated loudness from zap test – top 10 stations	54
Figure 5.27: Integrated station loudness comparison between zap testing and full testing.....	55
Figure 5.28: Overall loudness ranking of all 23 radio stations available at testing high-site.....	56
Figure 5.29: Loudness range as indicator of problematic fluctuations.....	57
Figure 7.1: Distribution of absolute error for evaluated loudness models	69
Figure 7.2: Loudness distributions of all adjacent programmes in radio stations	71

List of abbreviations and acronyms

ADC	Analogue to Digital Conversion
AL	Alignment Level
ARIB	Association of Radio Industries and Businesses
ATCS	Advanced Television Systems Committee
CALM Act	Commercial Advertisement Loudness Mitigation Act
CRC	Communications Research Centre Canada
DAB	Digital Audio Broadcast
DAC	Digital to Analogue Conversion
dB	Decibel
dBFS	Decibel Full Scale
dBm	Decibel with reference to 1 milliwatt
dBTP	Decibel True Peak
dBu	Decibel with reference to approximately 0.775 volts RMS
DR	Dynamic Range
DRT	Dynamic Range Tolerance
DSTV	Digital Satellite Television (Company)
DVB	Digital Video Broadcast
EBU	European Broadcasting Union
FIR	Finite Impulse Response
FM	Frequency Modulation
ICASA	Independent Communications Authority of South Africa
IEC	International Electrotechnical Commission
IP	Internet Protocol
ISO	International Organization for Standardization
ITU	International Telecommunication Union
L_{eq}	Equivalent Sound Level
$L_{eq}(RLB)$	Equivalent Sound Level with Revised Low-frequency B-weighting
LFE	Low-frequency Effects
LKFS	Loudness units, K-weighted relative to Full Scale
LRA	Loudness Range
LU	Loudness Units
LUFS	Loudness Units relative to Full Scale
ML	Measurement Level

NPC	Non Profit Company
PCM	Pulse Code Modulation
PML	Permitted Maximum Level
PPM	Programme Peak Meter
QoE	Quality of Experience
qPPM	quasi-Peak Programme Meter
RAB	Radio Advertising Bureau (NPC)
RAMS	Radio Audience Measurement Survey
RF	Radio Frequency
RMS	Root Mean Square
RSG	Radio Sonder Grense (Company)
SAARF	South African Audience Research Foundation (NPC)
SABC	South African Broadcasting Corporation
TAMS	Television Audience Measurement Survey
TPL	True Peak Level
TV	Television
VU	Volume Units

Terminology and concept clarification

Audio Compression

A compressor is an audio processor (analogue or digital) used to reduce the dynamic range of an audio signal by automatically applying varying levels of gain. (Not to be confused with data compression, used to reduce the size of data files.)

Audio Meter

A device that indicates audio levels visually. These can be hardware devices that display the measurements of analogue signals, or their digital equivalents that display the measurements from audio samples in digital audio.

Audio Normalisation

Audio normalisation refers to a uniform gain change of an audio signal to match a target level. With peak normalisation, the gain of the entire signal is raised until the highest audio peak is at a certain level, usually just below the theoretical maximum of digital audio (0dBFS) or other broadcast target levels. Audio normalisation does not change the dynamic range of the signal.

Dynamic Range

The dynamic range is a ratio, expressed in decibel, representing the range from loudest to softest signals that a system can handle without distortion (in the case of analogue audio), the theoretical range of values that can possibly be described (in the case of digital audio) and the range between the threshold of perception and threshold of pain in hearing (in the case of auditory perception).

Headroom

Headroom describes the amount of dynamic range available between the nominal operating audio level, and the onset of clipping or distortion.

Listener ‘Comfort Zone’

The ‘comfort zone’ describes the extent to which the perceived loudness can vary from one audio segment to another without causing remedial action by a typical viewer (i.e. changing the TV volume). It is measured as a relative change in Loudness Units (LU). The limits of the ‘comfort zone’ is given of +3LU and -6LU from one programme segment to the next.

Loudness Fluctuation

This is the term used throughout this dissertation for discrepancies in the perceived loudness of audio material when changing from one TV or radio station to the next, or when there is a change from one programme to the next (e.g. a commercial break after a film).

Loudness Unit

The Loudness Unit is used to differentiate from level measurements (usually expressed as decibel value). It is different in that the measure incorporates some psychoacoustic model to compensate for perceptual discrepancies in the way loudness is perceived. A relative change can be expressed in Loudness Units (LU), while absolute loudness levels are described as Loudness Units Full Scale (LUFS).

Peak Programme Meter

A type of audio metering device designed to display peak audio levels, predominantly for the purpose of avoiding overload distortion in digital systems. Peak programme meters have a fast response time (between 5-15ms), but are still unable to indicate the true peak of an audio signal.

Permitted Maximum Level

The Permitted Maximum Level refers to broadcast levelling standards that specify a specific level that audio signals may not pass. It is required to prevent the over-modulation of signals (or overloading of a transmitter), and audio meters are used to keep audio signals under the Permitted Maximum Level.

Volume Unit Meter

A type of audio metering device designed to display average audio levels, especially useful in the analogue domain where the audio reference level (e.g. +4dBu) describes the maximum permitted level of steady signals. VU meters are usually calibrated that their 0 marking aligns to the +4dBu reference level. This meter has a response time of approximately 300ms, and is unable to display transients (quick changes in level) accurately.

‘Zap Test’ Simulation

This describes the process of manually simulating a scenario where a viewer or listener would rapidly tune through a number of TV channels or radio stations with the purpose of selecting one. This process can lead to loudness fluctuations (described above).

Chapter 1

Introduction: loudness fluctuations in broadcast audio

This chapter includes background information to the study, the purpose and aims, and the research objectives. It explains the significance and motivation of the research problem, including basic assumptions, limitations and a delineation of the study. The chapter concludes with a basic overview of all the chapters in the dissertation.

1.1 Introduction and identification of the research theme

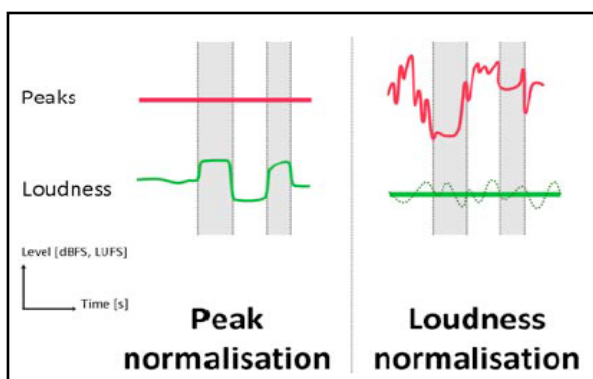
Consumers complain about loudness fluctuations in broadcast audio (Couling 1997; Riedmiller, Lyman & Robinson 2003; Spikofski & Klar 2004: 1), and it has “... become a major source of irritation for television and radio audiences around the world – that of the jump in audio levels at the breaks within programmes, between programmes and between channels” (Camerer 2010: 1). The problem is often linked specifically to commercials (advertisements), but it is a common misconception that broadcasters intentionally broadcast commercials louder than other programming to attract the attention of the audience (FACTS 2002: 3). Commercials *are* often louder, and it is to attract attention, but it is caused by high levels of fast acting compression applied by the content producers exploiting the peak level normalisation paradigm (Moore, Glasberg & Stone 2003: 1123; Miyasaka & Kamada 2006: 182). The problem is that South African analogue terrestrial broadcasts are currently only regulated by a maximum permissible peak level, and this level has little to do with the perceived loudness of programme audio.

Loudness is a subjective phenomenon. It is a “psychological term used to describe the magnitude of an auditory sensation” (Fletcher & Munson 1933: 377). It is the perceived strength of audio and is subject to a number of psychoacoustic considerations including duration, level, frequency and content of the audio. Loudness should not be confused with objective measures such as the level, amplitude, intensity or power of audio. Broadcasters and content producers have traditionally used level measurement and normalisation to set programming and broadcast levels, particularly by specifying a maximum permitted peak level. These measures are typically taken using either Programme Peak Meters (PPM) or Volume Unit (VU) meters. While controlling the maximum absolute peak level is necessary to avoid the electrical overloading of a broadcast transmitter, it is an arbitrary point relative to

the subjective loudness, and neither the PPM nor VU meter can represent perceived loudness or even the actual true peak level accurately (Skovenborg & Nielsen 2007: 2).

Modern dynamic range processing techniques have enabled producers to exploit the sub-distortion headroom below the maximum permitted level in a phenomenon described as the “loudness war”. Content producers and channels (especially in radio broadcasting) compete to be perceived as the loudest due to a simplistic belief that “louder is better” (Vickers 2010b: 1, 27). Other than loudness fluctuations, side effects of this “war” include reduced dynamic range, listener fatigue and an overall reduction in audio quality (EBU 2009: 1). “Undesired loudness changes in television have been around for a long time. The causes included an inconsistent use of available audio dynamic range, inadequate specifications and practices for loudness control, and the absence of a recognized and widely accepted method for accurately measuring the subjective loudness of audio signal” (Norcross, Lavoie & Thibault 2011: 12-13). Due to these factors, the international broadcast community is currently undergoing a fundamental audio levelling paradigm change to counteract loudness fluctuations and address the other associated problems; a change from peak normalisation to loudness normalisation (see Figure 1.1) with Japan, Australia, Europe and the United States of America (USA) adopting standards¹ recommending or requiring loudness measurement and normalisation (Camerer 2010: 1). “Broadcast levels can be set more consistently using a perceptually based measurement system; a loudness meter” (Lund 2006b: 4).

Figure 1.1: Peak-level normalisation vs. loudness normalisation of audio.



Taken from Camerer (2010: 1).

¹ Japan: ARIB TR-B32, Australia: FreeTV OP-59, Europe: EBU R128, USA: Calm Act and ATSC A/85.

South African free-to-air broadcasters still have a peak levelling paradigm using Peak Programme Meters almost exclusively², but could benefit from a paradigm shift to loudness normalisation. While there is extensive international scholarship on various aspects of loudness, including the loudness of broadcast audio, there has not been any South African research on broadcast audio loudness. There has also been no specific research on the loudness of South African broadcast audio, and this topic therefore does not duplicate any studies. The proposed study is up to date with recent research findings in a field that is currently still undergoing the paradigm shift as mentioned above, and the international regulations and recommendations have only been agreed upon in the last five years, and are still actively being refined.

The author created an experiment to measure and analyse loudness data of South African broadcast audio (both television and FM radio) to investigate if there are loudness discontinuities both within a single station between content, and between stations. If there are loudness fluctuations, it will be possible to quantify the nature and extent of the fluctuations, whether a listener/viewer would take action to counteract these fluctuations, and whether a shift to a loudness normalisation paradigm would counteract the problem. The following section will present the research question that forms the basis of the proposed research methodology.

1.2 Identification and demarcation of the research question

Given the context provided by the previous section, the primary research question of the research is: *Does South African free-to-air broadcast audio contain loudness fluctuations of a magnitude that exceeds previously determined limits of listening comfort?* The assumption in response is: South African broadcast audio programming, television and especially radio, contains loudness fluctuations great enough to fall outside the previously determined listening ‘comfort zone’.

This question and assumption give rise to subsidiary research questions and related assumptions. Firstly, *what is the extent of loudness fluctuations in South African broadcast audio, both within and between channels, in television and radio respectively, and do they exceed the previously determined limits of listening ‘comfort zone’?* The subsidiary

² See section 2.2.4

assumption is: Fluctuations exist both when changing from one channel to another, and between different types of content (e.g. between a drama series and a commercial advertisement), and the experiment will quantify the extent of the fluctuations.

The second subsidiary question is: *Does the broadcast audio programming contain a Loudness Range that is appropriate for the intended platform, listening environment and any fluctuations that exceed previously determined limits of Dynamic Range Tolerance of the listener?* The subsidiary assumption is: The Loudness Range of programme audio varies greatly, and at times may not be appropriate for the intended platform.

The study is demarcated as follows. The key variable is long term integrated programme audio loudness measured according to the European Broadcasting Union (EBU) R128 recommendation, with fluctuations thereof plotted with reference to the listening “comfort zone” (EBU 2011a). Additional variables include the True Peak Level and the Loudness Range. The geographical area under investigation is the Republic of South Africa, specifically including free-to-air terrestrial television and radio broadcasts. The time period under scrutiny is 2015, before migration to digital broadcasting.

The following section will provide additional conceptual delimitation and very briefly underline the purpose and value of the study.

1.3 Delimitation, purpose and value of the study

This section will first highlight what is outside the scope of the intended research before explaining the purpose and benefits of the study.

It is important to take note that this study does not aim to introduce any new or refine existing scientific methods for measuring programme loudness, nor does it propose a novel way to combat loudness fluctuations in broadcast audio. These fields contain much scholarship and praxis with many regulators and broadcasters around the world already undergoing a paradigm shift from peak normalisation to loudness normalisation. Instead, the purpose of this study is simply to investigate and test if South African broadcast audio contains loudness fluctuations, and if found to be present, to evaluate the nature and extent of the fluctuations, and whether it would cause annoyance to a listener or viewer and detract from their quality of experience. This research focus is positioned within the field of post-production and broadcast

audio, which is a sub-domain of music technology, which is a multi-faceted and interdisciplinary field.

Seeing as no such research has been conducted in the country to date, the key purpose will be to establish baseline data for South African broadcast audio loudness. This will benefit further studies on the topic. The results and recommendations of this study could also form the basis of a broadcast policy recommendation. In this regard, the timing of the study is advantageous as the public broadcaster is yet to make a transition from analogue to digital terrestrial broadcast (DOC 2014). A paradigm shift from peak to loudness normalisation has cost implications, not only in new equipment, but also in training staff (EBU 2011d: 40). Therefore, if such a shift is considered, it would be ideal if the timing coincides with a shift from analogue to digital broadcast as this will have similar types of, albeit much larger costs implications.

1.4 Structure of the research

This chapter has given a brief introduction to the topic, and has laid out the research question for the study. Chapter two presents the literature review, including the historical context, theoretical framework and methods of the study. Chapter three explains the methodology of the study, the sample selection, technical considerations and methodology of data collection, ethical and legal considerations, and various approaches of data analysis. Chapter four presents all the basic radio and television findings, focussing on overall station loudness and loudness ranges, including breakdowns according to content types. Chapter five uses the data from chapter four and evaluates the nature and extent of any loudness fluctuations measured. This chapter also presents ‘zap test’ results, and all additional observations. Chapter six presents a summary of all the findings from the previous two chapters, and also highlights proposed future research in the subject field. It also explores how production houses and broadcasters can employ loudness normalisation to prevent these problems, including the benefits and challenges of making such a paradigm shift.

Various appendices present tables and figures too lengthy or cumbersome for the main text, including audience measurements surveys and loudness distribution graphs.

Chapter 2

Theoretical considerations: audio measurement and perception

This chapter includes a literature review and will explain key concepts in audio description, psychoacoustics and audio perception, as well as loudness measurement. It gives broad context to the *status quo* of the process, metrics, and requirements of broadcast audio. It also gives an overview of the academic discourse that has led to the paradigm shift from peak normalisation to loudness normalisation in other countries and regions.

2.1 Audio levelling

This section describes the different ways that the concept ‘dynamics range’ can be interpreted, explores the limitations of various transmission methods and media, and explains why this necessitates metering and other audio levelling standards, especially for broadcast audio. The section concludes with a brief overview of how these levelling standards have been the major contributing factor to broadcast loudness fluctuation problems.

2.1.1 Dynamic range, and medium limitations

In order to synthesise, reproduce and amplify sound, or to capture (record) sound, it goes through various devices that include gain or conversion stages. All recorded sound, whether analogue or digital, has to be stored in or transmitted through a specific medium. These devices and media have finite operable or dynamic ranges, and audio levelling is required for the appropriate or optimal use of these devices or media. The dynamic range of sound is expressed as “...the ratio of the largest to the smallest intensity of sound that can be reliably transmitted or reproduced by a particular sound system” (OED 1993), or more simply “the difference between the noise floor and the onset of distortion” (Gallagher 2009: 62). This range is presented as a logarithmic scale and usually given as a single decibel (dB) value.

In the analogue domain, the dynamic range refers the range (between minimum and maximum) of electrical voltage that may exist in the system. In digital audio it describes the range between the theoretical noise floor (due to quantisation noise) and digital full scale, and is determined by the bit-depth of digital conversion, as indicated in Table 1. The human ear also has a dynamic range, often given as approximately 120dB (Huber & Runstein 2014: 64).

Table 1: Digital medium dynamic range in terms of full-scale signal to quantisation noise

Digital binary word length	16-bit	20-bit	24-bit
<i>Theoretical dynamic range</i>	<i>98dB</i>	<i>122dB</i>	<i>146dB</i>

Adapted from Spikofski & Klar (2004: 5).

Before the broadcast implications of various dynamic ranges can be contextualised, it is important to define specific language used for audio level descriptors, especially since some terms are ambiguous and others often used incorrectly interchangeably (Katz 2007: 166).

2.1.2 The language of audio level descriptors

The importance of using the correct audio terminology should not be underestimated. A number of terms are used synonymously, but have distinctly different meanings. For instance, ‘amplitude’ describes the pressure of sound waves, ‘power’ describes the total energy of the sound source (amplitude over time, or Watts), and ‘intensity’ is the power measured at a specific distance. ‘Level’ is also a measure of intensity, but is an abstract term and needs to be used with a point of reference, while lastly, the term ‘volume’ should be avoided as it “is an imprecise consumer term with no fixed definition” (Katz 2007: 166). A Root Mean Square (RMS) measurement describes the use of the quadratic mean as “a method of averaging levels which computes the equivalent power of material” (Katz 2007: 309). All of these terms describe objective measures that, while having a correlation to, do not equate to the subjective phenomenon of ‘loudness’. These terms should therefore be used with care.

Vickers (2010a: 5) points out that “the use of the term ‘dynamic range’ is [also] potentially confusing” as the term is used in a popular ‘Dynamic Range Meter’. This meter measures the micro-dynamics of a recording through the “average cumulative difference between peak and RMS over a specific period of time” with only the top 20% of levels being taken into account (PMF 2015), and Vickers contends that this is more closely related to the concept of a ‘Crest Factor’. The term ‘Crest Factor’ describes the “peak-to-average ratio, or a signal’s peak amplitude divided by its RMS value” (Vickers 2010a: 3) and is also expressed in a single dB value. Vickers (2010a: 5) instead suggests ‘dynamic spread’ while Skovenborg and Lund (2008) suggested ‘consistency’ and ‘density’. All of these terms were meant to describe various forms of dynamic range, but have since been superseded by the descriptors outlined in section 2.2.5.

From a technical point of view, broadcasters are concerned with the sense of ‘dynamic range’ described in the previous section. This ensures that broadcast audio levels fall within technical limitations. To achieve this, such levels need to be measured, regulated by standards to ensure interoperability, and processed using dynamic range manipulation techniques where appropriate and/or desirable. These topics are covered in the next three sub-sections.

2.1.3 Level measurement and metering

Programme audio generally consists of tones with dynamic variation, and therefore a dynamic (moving) meter is required for measuring real-time audio signals. However, audio level meters have an inherent trade-off between how quickly (and accurately) they can display peak levels, and the ease of observing a dynamically fluctuating meter display. Spikofski and Klar explain that “many different programme meters are in use at professional studios, with widely varying ballistical features”. The USA, Australia and France predominantly used VU meters while the rest of Europe and most of the rest of the world predominantly used various versions (see Figure 2.1) of PPM or qPPM meters (ITU 1988: 2; Spikofski & Klar 2004: 2-3).

VU meters are reasonably slow to respond to input stimuli, with an attack or integration time of ± 300 ms. While these meters are marginally better at showing an average signal level due to the slow integration time, they are unable to indicate fast peaks or transients and hence require additional systemic headroom; up to 18dB (ITU 1988: 2; Spikofski & Klar 2004: 2-4; Skovenborg & Nielsen 2007: 2).

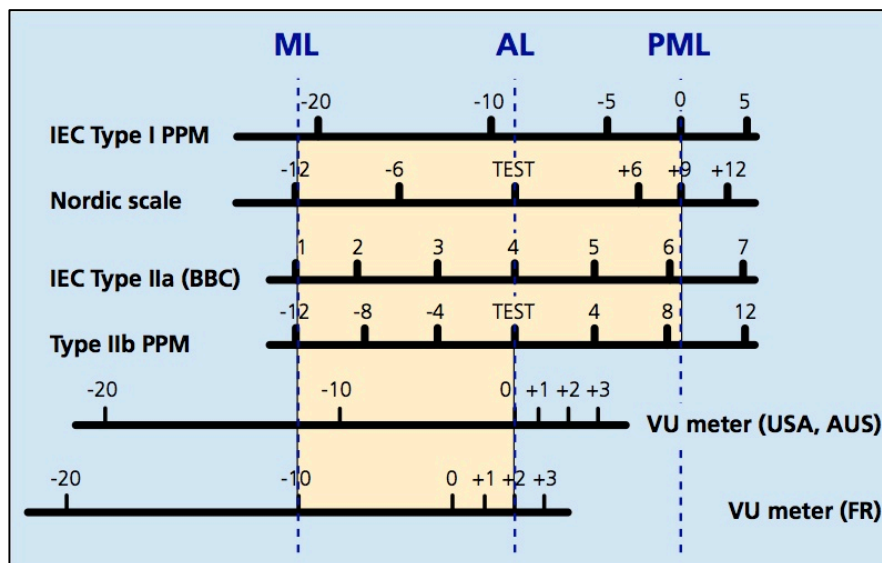
The technical standards for various types of PPM meters (as in Figure 2.1) were established in IEC Standard 60268 part 10 (IEC 1991). A PPM meter, as the name implies, is much better at indicating peak levels but still require at least 9dB of headroom due to an integration time between 5-10ms (1991: 29, 45). In digital audio, a fast PPM is technically able to achieve an integration time of 0ms, but even then is not able to indicate the true peak level the signal could achieve, due to its design as a sample meter rather than a signal meter, and its inability to detect inter-sample peaking (Lund 2006b).

Armed with standardised tools (meters) to measure the level of audio signals, broadcasters were able to introduce audio levelling standards based on these meters, as explained in the next section.

2.1.4 Broadcast audio levelling standards

Broadcast audio levelling standards specify a specific calibration, or Alignment Level (AL) that is 9 dB lower than the Permitted Maximum Level (PML) of the signal as shown in Figure 2.1 below. Some countries like the United Kingdom specified an AL of 8 dB below the PML to match their PPM scale. The measurement level (ML) was a further 12 dB below the AL and used for other equipment testing and measurement purposes (Emmett 2003: 1).

Figure 2.1: Programme alignment in the analogue world

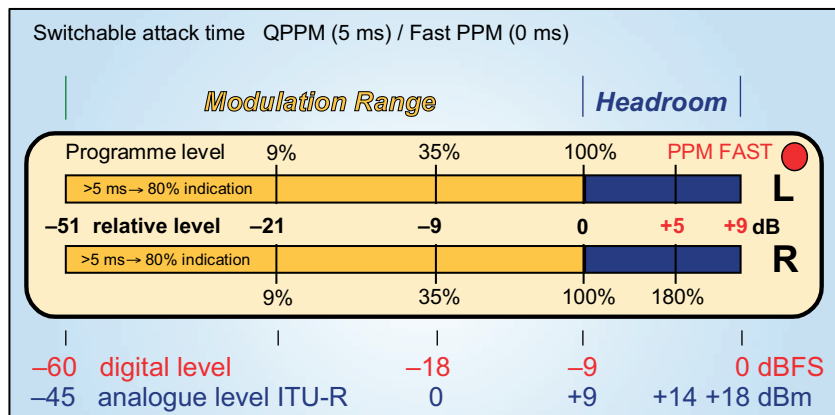


Taken from Emmett (2003: 1).

In analogue radio frequency (RF) broadcasts, measurements are taken relative to maximum permissible level of 100% modulation based on the regulations for the peak deviation frequency. With the widespread adoption of digital audio technology, the need for alignment levels to match between digital and analogue systems arose. The EBU recommendation R68 defines this relationship and suggests a further 9 dB of headroom, recommending an AL “18dB below the maximum possible coding level of the digital system” (EBU 2000: 1), in other words -18dBFS. This translates to -18dBFS = 0dBm for the AL as indicated in Figure 2.2. This -18dBFS alignment or calibration level is still found across the broadcast world³.

³ While the -18dBFS level is more common, the US digital PPM alignment level is defined at -20dBFS, also in the IEC 268-10 standard.

Figure 2.2: Broadcast peak programme meter



Adapted from Spikofski and Klar (2004: 3)

The South African public broadcaster, the SABC, uses a local variant of the Type IIb PPM with a 10ms response time and 24dB in 2.8sec return time. While most international broadcasters prescribed an AL 8 or 9dB below the PML, the SABC specifies an AL of only 6dB below the PML. The actual PPM would read -6 for the AL tone with a physical level of 0dBu, which would be 50% relative to a PML of +6dBu where the meter would read 0 (SABC 2003: 137; SABC 2004: 7). This dates back to before the introduction of EBU R68, but is still the current audio levelling standard for the SABC. Other South African broadcasters, including e.tv and Multichoice/DSTV both prescribed the use of a digital PPM with alignment at -18dBFS according to EBU R68 (Mahomed & Shabangu 2011: 16; Kruger 2013: 8).

While controlling the maximum absolute peak level is necessary to avoid the electrical overloading of a broadcast transmitter, it is an arbitrary point relative to the subjective loudness, and neither the PPM nor VU meter can represent perceived loudness or even the actual true peak level accurately (Lund 2006b: 3; Skovenborg & Nielsen 2007: 2). However, before the path of loudness measurement and the international standardisation thereof can be laid out, the next part will briefly describe how the peak levelling paradigm contributed to loudness fluctuations, and the so-called ‘Loudness War’.

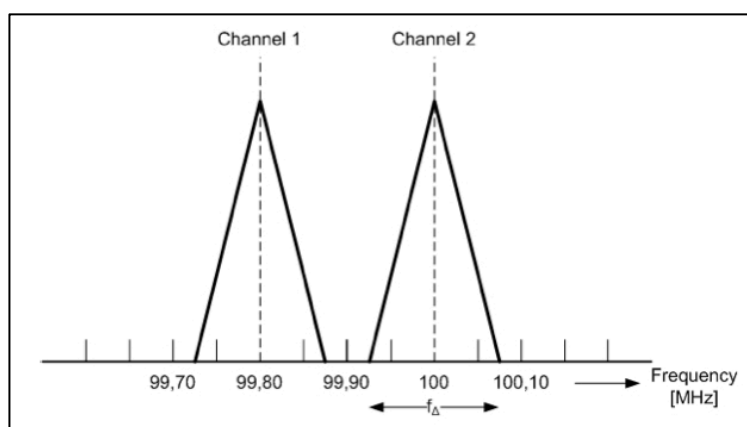
2.1.5 Dynamic range manipulation, and the ‘Loudness War’

The previous sections highlighted the necessity for appropriate audio levelling and dynamic range, and by extension, dynamic range manipulation for such purposes. This section will briefly explore why there is a link between the peak level normalisation paradigm and the “loudness war”.

In addition to using dynamic range manipulation for compliance, it is also used for artistic and unfortunately, for competitive purposes. The ‘loudness war’ or ‘loudness race’ describes a phenomenon, most clearly observed in popular music, where content producers compete with each other to be the loudest for the sake of being loudest, or due to a simplistic belief that ‘louder is better’ (Katz 2007: 187-188; Vickers 2010b; Weymouth 2012; Devine 2013; Taylor & Martens 2014). “Level control in digital audio production has traditionally been based on measuring peak level... Such practice makes material with low dynamic range appear louder, and has led to the so-called loudness war...” (Skovenborg & Lund 2008: 1).

Vickers explains that the term ‘loudness war’ has its origins in FM radio broadcasting in the late 1970s (2010b: 3). Terrestrial broadcast systems, such as those operating on an FM carrier is typically regulated based on the maximum deviation of the carrier, as illustrated in Figure 2.3, with a limiter to prevent the modulated signal and its sidebands from interfering with adjacent channels. As Camerer (2010: 1) explains, “[t]he carrier's maximum deviation ... was standardized in many countries at 50kHz and the PML at 30kHz deviation (equating to –9dBFS), which thus allowed 20 kHz, or 4.4 dB of headroom.” However, modern processing techniques enabled the manipulation of the dynamic range to such an extent that the PML could now be made equal to the full 50kHz frequency bandwidth, without much audible distortion, but with a loss of all transients. This provided a gain in loudness, but a loss in headroom (2010: 2). Radio stations exploited this to be louder than the next station, or not to be perceived as being ‘left behind’ by other stations becoming louder. Any station that does not exploit this and remains with the original standard, will be perceived as much softer, and hence cause an inter-channel loudness fluctuation when tuning between channels.

Figure 2.3: Frequency spacing of FM channels



Taken from Sack (2011: 6).

Vickers (2010b) and Weymouth (2012) provide fascinating analysis by looking at the ‘loudness war’ through the lenses of game theory, where players (radio stations) treat the race to louder broadcasting as a zero-sum game. Vickers gives an overview of similar ‘wars’ in movies, vinyl records and especially compact disks (or digital music recording in general). Some conclusions include that, while with two otherwise identical signals, the louder one is in fact perceived as better, the same does not apply with different signals, and even with the original signal, there is a definite ceiling in perceived utility. Furthermore, no correlation was found between louder music recordings and greater commercial success. Lastly, the hyper-compression and other processing tactics employed to increase loudness can severely affect the aesthetic appeal of the content (e.g. lack of dynamic range and loss of transient frequencies) as well as have noticeable and fatiguing side-effects through various types of distortion.

Thomas Lund contends that in television audio, the incentives and therefore behaviour of the listener differs to that of commercial music. “Please remember that louder is not better. Consistency is what counts, while sounds trying to grab our attention by being loud feel obtrusive and get deselected. If the listener wants it louder, she will turn up the volume control...” (Lund 2006b: 10). Television and radio broadcasters also contain various genres and material lengths. It is easier to hyper-compress and limit the sound of a 30 second long commercial compared to a feature film. If the latter is adjacent to (or inserted into) the former, loudness fluctuations will once again occur.

The inherent shortcomings of the peak levelling paradigm (leading to intentional loudness exploitation through dynamic range processing) is not the only cause of loudness fluctuations. Spikofski and Klar (2004: 1) contend that loudness fluctuations had also been caused, or was at least aggravated by the inexperience of channels, the use of various different level meters, the presence of archive material, and the lack of a standardised loudness meter up until that point.

2.1.6 Conclusion

This section defined important language needed to describe audio measurement and levelling, and explained why audio content needs to be levelled to be appropriate for the dynamic range of a storage medium or transmission method. It gave a brief overview of broadcast audio levelling standards, and highlighted why the peak levelling paradigm created by these standards was the major contributing factor to loudness fluctuations through the so-called ‘loudness war’.

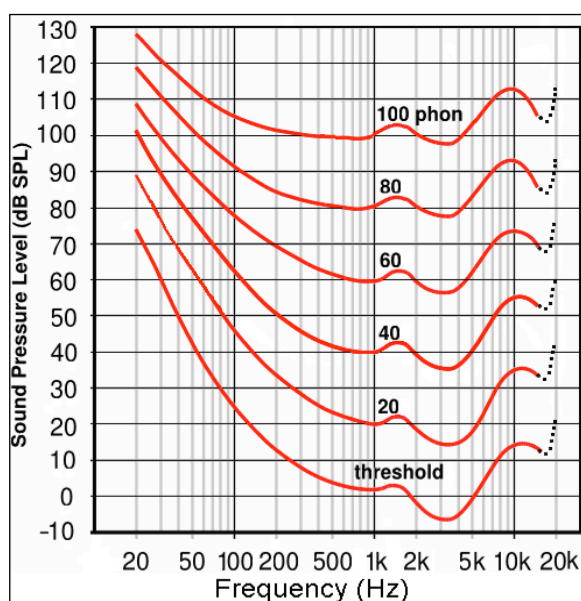
2.2 Measuring loudness

This section will first explore the challenges of measuring the subjective phenomenon of loudness by explaining the psychoacoustics of loudness perception, before giving an overview of various attempts of measuring loudness in the last century. It will also provide and justify the three main components of the theoretical framework of this study, namely a method for measuring audio programme loudness parameters, and two metrics for listener tolerance.

2.2.1 Psychoacoustics and loudness perception

The field of psychoacoustics is an intersection between psychological and physiological responses to sound and much research explores various non-linear aspects of human hearing. Some of the earliest work was published by Fletcher and Munson (1933) where they established differences in loudness perception based on the frequency of the sound, and they described this non-linear frequency response of the ear using a curve. They established that the ear was least sensitive for very low and very high frequencies, and that this response curve also differs according to the level of reproduction of the sound. Their original work was re-evaluated in 1956 (Robinson & Dadson 1956) and formalised by the International Organization for Standardization (ISO) in 1987 in their ISO226 standard using the name ‘equal-loudness contours’, which in turn was revised to a second edition in 2003 (ISO 2003), as presented in Figure 2.4.

Figure 2.4: Equal-loudness contours according to ISO226:2003



Taken from Sengpiel Audio (2005).

When a pure tone is reproduced at varying dB SPL according to a particular equal-loudness contour, the subjective loudness remains constant, and that contour represents a particular *phon* level. The *phon* scale corresponds to sound pressure level (dB SPL) at 1000Hz (i.e. at 1000Hz, 40 phon = 40dB SPL). These equal-loudness contours explain one of the most important reasons why VU and PPM meters cannot indicate loudness, as they cannot take frequency into consideration. The *sone* scale is based on the *phon* scale, but provides a relative (ratio) rather than absolute level. It is important to note that *phon* or *sone* scales would not be appropriate as units for this study as they are only accurate for continuous, frontal, steady pure tones (sine waves) while “real-world signals...are typically broadband and fluctuating” (Skovenborg & Nielsen 2004b: 4). These scales do not take into consideration any temporal or spatial aspects of loudness perception and even in terms of spectral aspects, only take a single, discrete frequency in consideration and not the effect of the total bandwidth.

An excellent summary of the aspects affecting loudness perception is presented by Skovenborg and Nielsen (2004b: 2-4). These include additional spectral aspects like *spreading* (how various frequencies interact when perceived together), and also the *spectral loudness summation* and *bandwidth* of the sound (how the bandwidth of the sound, wider than a pure tone, affects loudness perception). The ear does not perceive sounds shorter than approximately 100ms as loudly as exactly the same sound longer than 100ms (or a continuous version thereof). Other temporal aspects include acoustic *masking* and *temporal threshold shifts* in hearing. Lastly, loudness perception is also affected by the localisation of sound (where the sound is coming from) and spatialisation (the reverberant qualities) of the listening environment. Another spatial aspect, known as *binaural loudness summation*, takes into account the affect of the predominant mode listening, that of using both ears at the same time.

The spectral and temporal aspects of perception are also of particular importance in understanding why the peak normalisation paradigm causes loudness fluctuations. Put simply, two sounds at different frequencies but at the same amplitude or level will read the same on a VU or PPM meter, but be perceived differently in terms of loudness. Similarly, a very short and a longer sustained sound with the same level will register the same on a PPM meter, but again be perceived differently in terms of loudness. However, this does not imply that content producers and broadcasters have not been aware of the shortcomings of their meters for a long time; they had just lacked accepted methods to deal with the problem (Norcross, Lavoie &

Thibault 2011: 4). The following section gives an overview of the process that resulted in international standardisation.

2.2.2 International standardisation of loudness measurement

A loudness model attempts to predict the perceived loudness of an audio signal, and is implemented in loudness meters and loudness control devices (Skovenborg & Nielsen 2004b: 1-2). Many authors have published detailed overviews and subjective listening experiments to test various loudness measurement methods (Benjamin 2002; Moore, Glasberg & Stone 2003; Soulodre 2004; Skovenborg & Nielsen 2004a; Skovenborg & Nielsen 2004b), but up to that point there had been no wide acceptance or standardisation. Skovenborg and Nielsen (2004b) should especially be consulted for thorough analysis of various loudness measurement methods.

“A number of loudness measurement algorithms and methods have been developed over the years. The most popular ones includes Zwicker’s loudness model (ISO 532-1975), the CBS Loudness Summation Method, $Leq(m)$, $Leq(A)$ & $Leq(C)$. None of these techniques have been standardized nor widely accepted for use in the broadcast industry” (Norcross, Lavoie & Thibault 2011: 4).

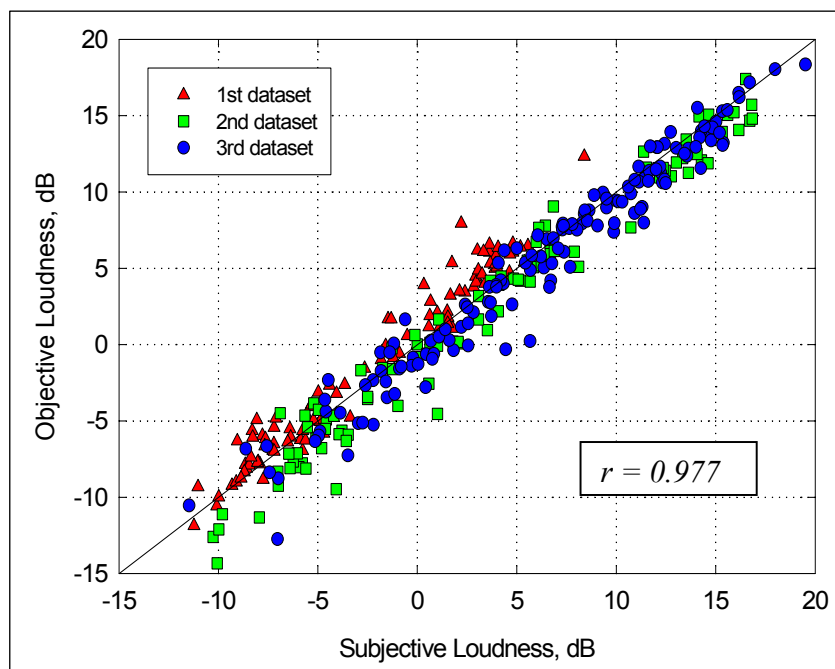
There has been some success with metadata based schemes such as ReplayGain in pop music, and the dialnorm and prog_ref_level tags in cinema, but these solutions have not been standardised across media, and have been found not always being robust, especially when metadata is set incorrectly or not set at all (Norcross, Lavoie & Thibault 2011: 5; Camerer *et al.* 2012: 12).

The International Telecommunication Union (ITU) formed a Special Rapporteur Group in 2000 to find an objective loudness measurement method for measuring broadcast audio, amongst other objectives. The process involved multiple independent subjective listening tests, and eventually resulted in the publication of the ITU-R BS.1770 recommendation entitled *“Algorithms to measure audio programme loudness and true-peak audio level”* in 2006, with revisions in 2007, 2011 and 2012 (ITU 2012).

A summary and evaluation of submissions to the group was published by Soulodre (2004) and also examined by Skovenborg and Nielsen in an extensive study to evaluate a range of different loudness models (2004b) as well as listening test methods and the statistical evaluation of

results (2004a) (see Appendix A). The $L_{eq}(RLB)$ loudness model was selected and extended for multichannel audio, and “the new multichannel algorithm retains the very low computational complexity of the monophonic $L_{eq}(RLB)$ algorithm” (ITU 2012: 8) as given in Figure 2.6. This algorithm is simple to implement as it consists of two filters, a positive high shelf filter at about 4kHz at +4dB, and a modified second order high pass filter. This was named the K-filter, but should not be confused with Bob Katz’s K-system presented in his book (Katz 2007). The algorithm then takes a mean square of the filtered signal, adjusts for localisation (the rear channels are biased by approximately +1.5dB to account for psychoacoustic effects) and the channels are then summed (Figure 2.6). A loudness meter based on ITU-R BS.1770 was tested in five separate locations and showed a very high correlation between subjective and objective loudness, as evident in Figure 2.5.

Figure 2.5: A comparison of objective and subjective loudness for the ITU loudness meter

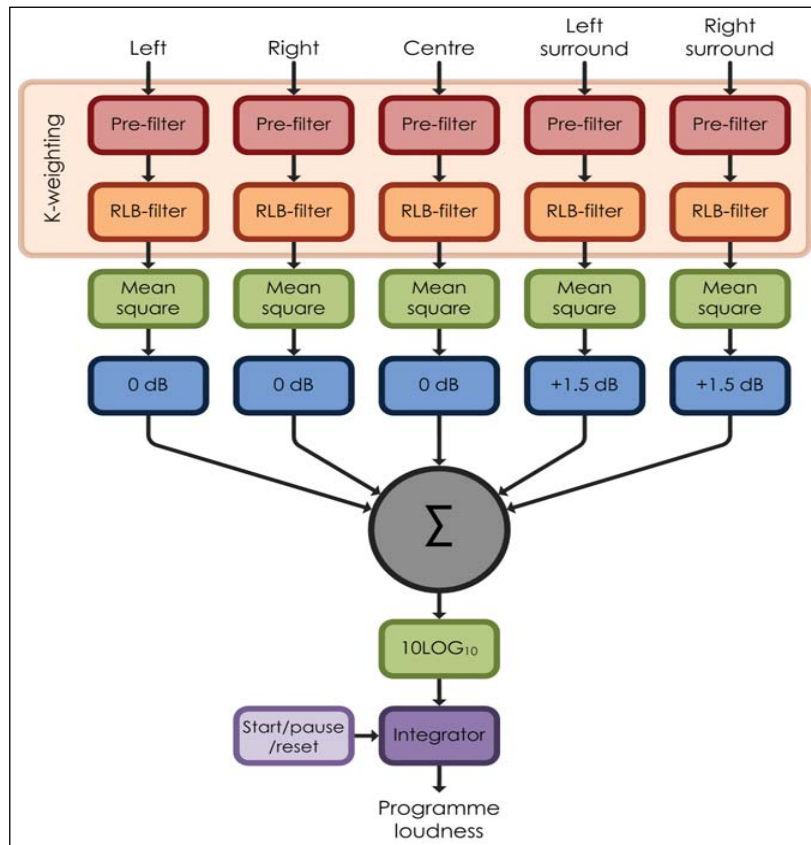


Taken from Norcross, Lavioe and Thibault (2011: 7).

ITU-R BS.1770 describes methods to measure the loudness and true peak of a signal, but does not define any prescribed values for content producers or broadcasters. That has been the work of regional organisations, and while their requirements vary slightly, they are all based on the ITU-R BS.1770 method, including the following international loudness standards: ARIB TR-B32 in Japan, OP-59 in Australia, EBU R128 in Europe, and CALM Act and ATSC A/85 in North America. To differentiate loudness from level measurements, the results

are expressed in Loudness Units (LU), but whilst retaining a direct proportion to decibel values, for the sake of equivalent relative changes.

Figure 2.6: Channel processing and summation in ITU-R BS.1770



Taken from Camerer (2010: 3).

Amongst the regional organisations, the PLOUD group at the European Broadcasting Union (EBU) developed the R128 recommendation (EBU 2011a), prescribing specific measurement windows and values for integrated loudness (a single value that gives the average across the entire programme), true peak, and introducing a new metric known as programme loudness range (LRA) (EBU 2011b: 5-8). The single target integrated loudness value is -23LUFS, and the maximum permitted true peak value is -1dBTP, but can still depend on encoding types (EBU 2011a: 4). The PLOUD group included two gates in the measurement method to prevent silence or very low levels from skewing the loudness results. The first is an absolute audio measurement gate at -70LUFS, and the second is a relative audio measurement gate at -8LU relative to the integrated loudness value of the programme. These were included in the third and latest revision of ITU-R BS.1770-3, however a relative audio measurement gate of

-10LU was employed despite research pointing to better results with a -8LU relative audio measurement gate (Grimm, Skovenborg & Spikofski 2010; Lund 2011).

Having a method for measuring integrated loudness that includes a target value allows for loudness normalisation, the details of which will be explained in the next section.

2.2.3 Universal loudness descriptors and EBU R128

The target value of -23LUFS integrated loudness is the single most important recommendation and metric for loudness normalisation. However, Skovenborg and Lund explain, “a single loudness descriptor is not sufficient to characterize different program genres and mixing-styles. Three different properties, relevant to the audio engineer, should be monitored and controlled independently” (2008: 2). This section will explain the need for, and difference of these three properties, and the relevance they have to the methodology of this study. These properties are *integrated loudness*, *true peak-level* and *loudness range*.

Integrated loudness describes the average loudness of an entire programme, and the measurement is robust in such a way that it can be used for a short interstitial under ten seconds, as well as a full feature length movie. The target value of -23LUFS allows for sufficient headroom in any genre. There is no difference between the units LUFS and LKFS and they can be used interchangeably. They refer to an absolute loudness relative to full scale, while LU is a relative measure (Camerer 2010; EBU 2011a; Lund 2011).

True peak-level is not dissimilar to the PML of analogue broadcast, as it is prescribed to prevent transmission over-modulation or digital clipping. The general prescription according to the EBU R128 recommendation is -1dBTP, as measured by an ITU-R BS.1770 true peak meter. A true peak meter differs from a PPM or other peak meters by employing four times oversampling in metering, enabling it to reveal inter-sample peaking. It is also advised that a maximum of -3dBTP is adhered to for any data reduction codecs (Lund 2006b; EBU 2011a; EBU 2011b).

Loudness range or LRA is the third property, but unlike the previous two is descriptive only and does not carry any prescribed targets. It is explored in further detail in section 2.3 on listener tolerance (EBU 2011c). The EBU R128 recommendation also makes provision for two additional, shorter ‘views’ of the loudness measure, known as *short term loudness* and *momentary loudness*. These are also not prescriptive, but are built into all EBU compliant

loudness meters, and their ballistics are useful while mixing audio. As explained by Lindström (2013: 8-9), these shorter views simulate temporal loudness perception, and are created by adding a smoothing filter after the summation phase indicated in Figure 2.6. This smoothing filter is achieved by “*rectangular integration windowing, also known as Finite Impulse Response, FIR, or square sliding-window RMS*” (Lindström 2013: 9). *Short term loudness* is an ungated 3 second sliding rectangular window measure, while *momentary loudness* is an ungated 0.4 second sliding rectangular window measure (EBU 2011b: 5-8).

It is important to note that there can never be a perfect loudness meter. Various subjective aspects of loudness perception, such as the same person perceiving the loudness of the same stimulus differently on different days (and many other aspects) will make this phenomenon inherently subjective. However, the methods and meters mentioned in the previous sections are not only better than we have ever had before, they have been adopted internationally, and much research is still going into their improvement.

2.2.4 Conclusion: South African adoption of loudness normalisation

To date, the SABC, e.tv and none of the local radio stations have implemented, or indicated the intention of implementing any form of loudness normalisation. Delivery specifications still make exclusive reference to PML and PPM values (SABC 2003; SABC 2004; Kruger 2013).

At the time of writing, Multichoice / DSTV is the only South African operator or broadcaster that had implemented any type of loudness normalisation. They selected the EBU R128 standard, and it is mentioned as early as 2011 (Mahomed & Shabangu 2011: 2), with a requirement for all short-form content to be R128 compliant coming into force on 14 August 2014 (DSTV 2014). They have also started the transition for other locally delivered content, and conduct extensive loudness monitoring, however, they are not a free-to-air broadcaster.

2.3 Listener tolerance

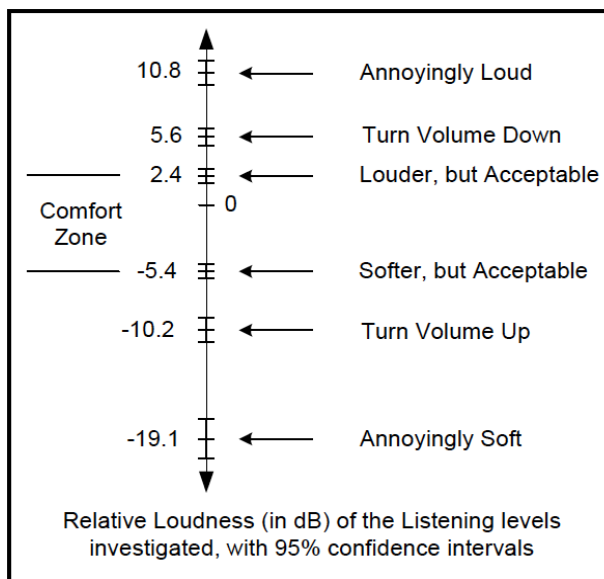
While universal loudness descriptors were defined and their adoption in open international standards explained in the previous section, the theoretical framework of this dissertation still requires metrics for gauging if, and how, a listener would tolerate fluctuations in these programme loudness descriptors. The first metric describes the ‘comfort zone’ of a listener, and the second describes the dynamic range tolerance. These metrics are the basis of both

subsidiary research questions of the dissertation. It is important to note, however, that fluctuations in loudness is not the only factor that affects the annoyance of the listener, and the effect of tonal, impulsive and other components also play a role (Skovenborg & Nielsen 2004b: 6). These factors are beyond the scope of this dissertation.

2.3.1 Listening 'comfort zone'

Measuring the programme loudness will indicate the presence and extent of loudness fluctuations, but an additional theoretical framework is required to predict how an audio consumer would react to these fluctuations. This is known as the loudness 'comfort zone', a concept first introduced by Riedmiller, Lyman and Robinson (2003: 3-7) who found that a listener will turn down the volume on their device if the programme loudness increases by 5.6dB and similarly, turn it up with a decrease of 10.2dB or more (see Figure 2.7).

Figure 2.7: Listener loudness 'comfort zone'.



Taken from Riedmiller, Lyman & Robinson (2003: 6).

However, the methods of this study predated the ITU-R BS.1770 standard, and included a measurement algorithm that did not measure anything other than spoken voice. In response, Skovenborg and Lund (2009) refined this concept in accordance with the ITU-R BS.1770 recommendation, and found that 50% of an audience will adjust the volume of their device for a loudness increase of 3 Loudness Units (LU) or a loudness decrease of 6 LU; and 95% of an audience will adjust the volume for a loudness increase of 5 LU or a decrease of 8 LU as given in Table 2 below.

Table 2: The effect of inter-program loudness fluctuations on the audience

	50% of subjects would adjust	95% of subjects would adjust
Loudness increase between programs	3 LU	5 LU
Loudness decrease between programs	6 LU	8 LU

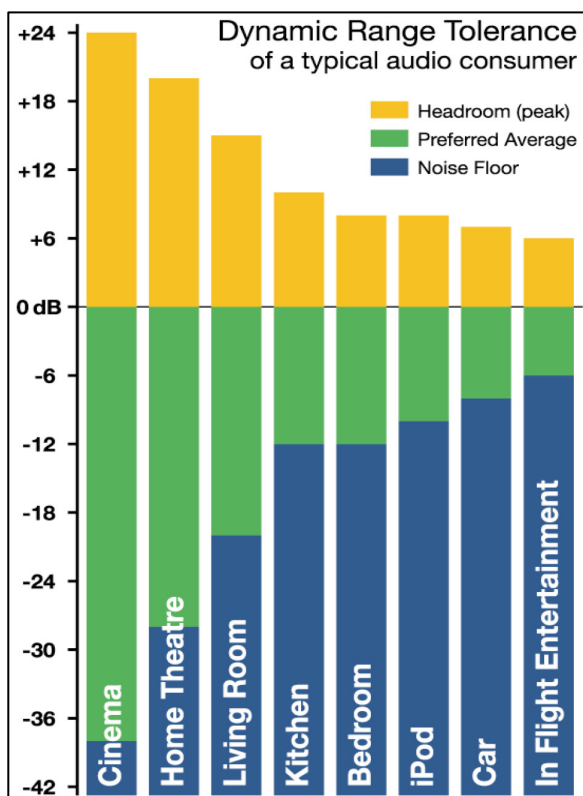
Reproduced from Lund (2011: 4).

This forms the key metric for evaluating listener annoyance or action due to loudness fluctuations and it is an important contributing factor to the overall Quality of Experience (QoE) of the broadcast (Kuipers *et al.* 2010: 219).

2.3.2 Dynamic range tolerance

An additional metric is required to analyse the suitability of programme audio with regards to the Dynamic Range Tolerance (DRT) of a listener (see Figure 2.8).

Figure 2.8: Dynamic Range Tolerance for various listening situations.



Taken from Lund (2006a: 57).

This is dependent on the location and type of device for which programme audio is intended, from here on referred to as the platform. This metric depends on the measurement of the LRA according to EBU R128. It is important to note that LRA is, in effect, a measure of macro-dynamics of an entire audio programme, especially long form content. It is the difference between the average loudest parts and average softest parts of an entire programme, and should not be confused with micro-dynamic measures such as the Dynamic Range (DR) or crest factor (explained in section 2.1.1). The LRA is based on short term values in 3 second intervals, and hence, due to sampling insufficiency, does not give a meaningful measure for programmes shorter than 30 seconds (EBU 2011d: 42).

It is important to keep in mind that both metrics mentioned in this section use the loudness data gathered through the loudness measurement method mentioned. The next and final section will briefly list a number of scenarios where loudness fluctuations could be encountered, and how this will be quantified by the theoretical framework from this chapter.

2.3.3 Conclusion: loudness fluctuation scenarios

Specific scenarios where a viewer or listener may experience loudness fluctuations have thus far only been alluded to. The most common example cited is advertisements appearing louder than adjacent programming (FACTS 2002; Moore, Glasberg & Stone 2003; Miyasaka & Kamada 2006; EBU 2011d: 42). This includes any type of short-form content. Short-form content is not strictly defined, but usually refers to items 30 seconds and shorter, and includes advertisements (commercials), trailers, bumpers, wraparound segments, promos (promotional programmes) and any other broadcast interstitials. The loudness fluctuations occur when they are inserted into another programme (e.g. a commercial in the middle of a feature film), or when adjacent to another programme. This is the first type of fluctuation that will be investigated, and corresponds to the type or genre of the content. However, there are other scenarios with the possibility of loudness fluctuations that will also be investigated.

Adjacent content (regardless of length) can produce fluctuations. When switching between television channels or tuning through various radio stations, fluctuations can occur from one channel or station to the next. Lastly, when content moves from a small loudness range to a large loudness range, this may also be considered a loudness fluctuation even though the average loudness may have remained constant. This fluctuation may still cause listener annoyance if the loudness range is larger than what would be appropriate for the platform. A

mixture of wide and narrow loudness range material will also increase the statistical probability of loudness fluctuations between material (Skovenborg & Lund 2009).

2.4 Conclusion

The second chapter has served as a review of the most relevant literature, while laying out the theoretical framework for the study. It has defined the language needed to differentiate between level and loudness and explained how the peak levelling paradigm, although necessary for technical compliance, has been one of the leading contributing factors to loudness fluctuations in broadcast audio. The non-linear nature of human hearing was explored, especially those aspects that affect loudness perception - an inherently subjective phenomenon. The chapter also explored the advances made in achieving an objective measure, and by extension a loudness meter (despite the inherent subjectivity) culminating in the international standardisation and adoption of a loudness measuring algorithm called the ITU-R BS.1770, including the EBU R128 implementation, specifically developed for broadcast audio.

Chapter 3

Experimental methodology

This chapter will introduce the research design and explain the methodological aspects of the research instruments used in the experiments. It will cover the equipment used, including important calibration considerations. It will explain the data output in terms of the key metrics of loudness measurement, and describe the statistical analysis employed.

3.1 Technical considerations

The research will first and foremost be a laboratory study or designed experiment to collect primary data (Mouton 2001: 155-156). This section will first discuss technical considerations of the experimental design of real-time loudness data measurement logging procedures, including equipment and calibration, and sample selection.

3.1.1 Hardware

As suggested by Riedmiller, Lyman and Robinson (2003: 9) “periodically logging the short-term ... measurement values against time can be used to identify and correct endemic errors in broadcast. This simple, yet very effective means of analysis can clearly show whether a channel has the proper level ... or has a large loudness variation during or between programs.” This is achieved by connecting both an FM radio and television receiver (Yamaha RX-V659) to an Analogue to Digital Converter (ADC), which in turn is connected to a computer running a real time programme loudness meter/analyser. The author used a Focusrite Saffire Liquid 56 Firewire audio interface as ADC.

3.1.2 Loudness metering and logging

Loudness metering was done with two meters. First, the free Orban loudness meter capable of exporting comma delimited data files of loudness data for further analysis. This meter is fully compliant with ITU-R BS.1770 and EBU R128 (Orban 2014). The data files include a date and time stamp that are collated with programme scheduling data freely available from broadcasters. For further redundancy and real-time monitoring abilities, the incoming signal was also routed into Steinberg Nuendo version 6.5 digital audio workstation that also has compliant EBU R128 loudness features (Steinberg GmbH 2015). The built-in Nuendo control

room loudness meter was used as a second meter, and loudness history graphs were generated using Nuendo loudness tracks.

3.1.3 Calibration

Equipment was calibrated according to the method described by the EBU (2011b: 8-9). Calibration signals will be sent through the signal path used for loudness analysis. The hardware described in the experimental design was selected because of the option for setting or verifying input gain digitally. This allows the input side gain staging to be kept consistent during the entire experiment. Radio frequency (RF) signals need to be strong for proper data logging, and each station and channel was checked manually to avoid multipath distortion or other possible causes for a weak signal. Terrestrial RF broadcasts are measured relative to maximum permissible level of 100% modulation based on the regulations for the peak deviation frequency. However, the purpose of the study is not to compare the output levels measured at the receiver end to absolute calibrated levels on the broadcasters side, but rather to compare the relative levels of various stations and programmes on the receiver side. Therefore, consistency in the calibration in this experiment is the most important part of the experimental design rather than trying to match the type of calibration used by various broadcasters, which may differ.

3.2 Sample selection

This section will describe the process of sample selection, first explaining which television channels and radio stations were selected and why, before highlighting which samples were captured from these broadcasters.

3.2.1 Channel and station selection

Loudness data will be recorded from all free-to-air television channels namely SABC 1, SABC 2, SABC 3 and e.tv. Recording data from every single FM radio station would not be feasible in terms of time and geographical availability. Instead, the ten most popular radio stations according to the South African Audience Research Foundation (SAARF 2014) Radio Audience Measurement Survey (RAMS) were selected for comprehensive data capture, as presented in Table 3. This table only presents the top ten; please refer to Appendix B for the full survey. The average given represents the percentage of the total South African population in all nine provinces for adults over 15 listening to that particular station. Full RAMS methodology available at SAARF (2015b).

Table 3: National Radio Audience Measurement Survey for May 2014 - June 2015: Top 10

STATION	MAY '14	JUNE '14	AUG '14	OCT '14	DEC '14	FEB '15	JUN '15	Average	National Rank
UKHOZI FM	20.2	20.5	20.5	20.9	20.7	20.5	20	20.5%	1
METROFM	17.3	17	16.9	18	17.4	17	16.9	17.2%	2
UHLOBO WENENE FM	11.9	12	12.1	11.7	11.4	11.3	12.3	11.8%	3
LESEDI FM	10.9	10.7	10.5	10	9.6	9.7	9.6	10.1%	4
THOBELA FM	8.8	9	8.9	9	9	9.2	7.9	8.8%	5
MOTSWEDING FM	9.1	8.9	8.8	8.2	8.3	8.3	8	8.5%	6
JACARANDA 94.2	5.5	5.6	5.5	5.1	5.1	5.1	5.2	5.3%	7
RSG	5.1	5.1	5.1	5.1	5.2	5.1	4.9	5.1%	8
GAGASI FM*	5.1	5	5	5	5	5	4.4	4.9%	9
5FM	5.2	5	5.3	5.1	4.6	4.5	4.6	4.9%	9
IKWEKWEZI FM	4.8	5	4.7	4.8	4.9	4.7	4.5	4.8%	10

* While Gagasi FM scored joint 9th place for the period under investigation, it is only broadcast in Kwa-Zulu Natal and would therefore neither be nationally representative, nor geographically feasible for data collection.

Adapted from SAARF (2015a: 5).

As noted in Table 3, Gagasi FM was ranked 9th, however is only broadcast in Kwa-Zulu Natal. This station was not included in the study, firstly because it would not be nationally representative (as the other stations in the top ten), but also as it would be geographically unfeasible for data collection within the scope of the study. The 11th ranked station, Ikwewezi FM, was included instead. In addition to the top ten stations, limited data was also captured for four other stations:

- *Talk Radio 702* as it is predominantly a talk radio station
- *Classic FM* as it is the only station available at the test site broadcasting classical music
- *Power FM* as it is one of the newest entrants on the commercial radio market
- *SAFM* as one of the older commercial radio stations

As evident from the overall listener percentage averages presented in Table 3, this sample provides the most comprehensive coverage of adult South African radio listeners for the least amount of radio stations evaluated. Additional methods for attaining comprehensive coverage is described in section 3.2.3.

3.2.2 Time and genre sample selection

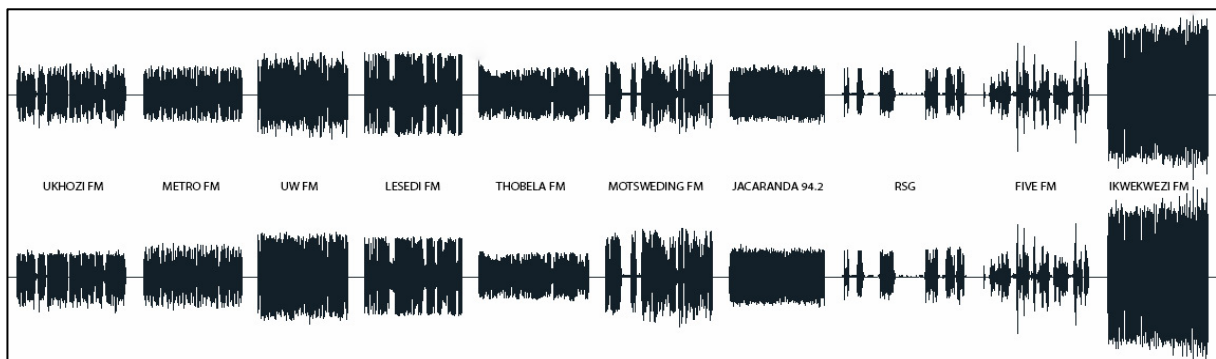
Samples are taken to include full listening time regions including breakfast, daytime, drive time, prime time and late night sections, and will therefore include all main broadcast genres.

For radio, data is collected in segments with multiples of 20 minutes (the longest continuous segments being 80 minutes). This method is in line with other similar studies (Riedmiller, Lyman & Robinson 2003; Miyasaka & Kamada 2006). These segments all start approximately two minutes before the hour in order to capture adjacent audio blocks that usually include a change on the hour. Approximately 100 programme segments are captured per station. A combination of schedule consultation and signal monitoring ensures that all possible genres are captured on all stations, depending on the type of station programming. For television data, full days of data is collected over the weekend and mid-week, which is feasible since only 4 channels are investigated.

3.2.3 Radio Zap Tests Simulations

A ‘zap test’ is a method proposed by Skovenborg and Lund (2009) to test normalisation schemes against loudness fluctuations that occur when ‘zapping’ (rapidly scanning) through channels or stations. In addition to the fourteen stations for which extensive data was collected, a number of ‘zap test’ simulations (see Figure 3.1) were also performed for all FM radio stations with strong reception at the test location in Rayton, Gauteng. This location was chosen as it was a radio high site, with strong reception of both Pretoria and Johannesburg based transmitters, as evident from the station duplication in Table 4. These tests simulate a real life scenario that occurs whenever a listener tunes (manually or automatically) from station to station. The tests were conducted with the radio tuner automatically tuning from station to station, and loudness data was logged for 5-10 seconds per station ‘zap’. The full list and order of stations are given in Table 4.

Figure 3.1: Example single “zap test” simulation of the top 10 radio stations



Although these tests were performed to fall within the same time and genre slots as mentioned in the previous section, it does not allow for genre specific intra-station loudness analysis as loudness data for full programme segments are not captured. In addition to simulating a real

life scenario of rapid scanning through stations, the secondary purpose of performing thorough ‘zap tests’ is to test whether this would be a more efficient method for monitoring the station loudness of a higher number of stations. The level of robustness will be verified by comparing the stations loudness results of the main fourteen stations in overall testing, to their station loudness results based on their ‘zap test’ segments only.

Table 4: Order of FM radio stations according to frequency for ‘zap test’ simulation

Order	MHz	Station	Order	MHz	Station
1	87.9	Thobela FM	16	97.5	Radio 2000
2	88.4	Lesedi FM	17	89.0	FIVE FM
3	89.3	Ligwalagwala FM	18	98.7	Power FM
4	89.3	Motsweding FM	19	99.2	YFM
5	90.1	Thobela FM	20	100.1	Phalaphala FM
6	91.0	Motsweding FM	21	101.0	RSG
7	91.5	Ukhozi FM	22	102.4	Ukhozi FM
8	92.4	Metro FM	23	102.7	Classic FM
9	92.7	Talk Radio 702	24	104.2	Radio Pretoria
10	93.2	UW FM	25	104.6	Safm
11	94.2	Jacaranda FM	26	105.7	Safm
12	94.7	Highveld 94.7	27	106.0	Talk Radio 702
13	95.6	Munghana Monene FM	28	106.3	Ikwekwezi FM
14	96.4	Metro FM	29	106.8	Lotus FM
15	96.8	Ikwekwezi FM			

3.3 Ethical and legal considerations

The research has been designed to avoid ethical or legal problems. The first and foremost legal consideration when including broadcast content in academic research is intellectual property rights and copyright of the programme material. By design, only loudness measurement data will be kept and archived, therefore eliminating the threat of infringing these rights. In addition to the loudness data, only basic other programme and broadcast metadata will be collected, including channel/station name, time of broadcast and genre (and other similar content descriptors).

3.4 Data analysis

Macro level analyses will compare fluctuations between the long-term integrated loudness values of stations, between listening time regions across stations and between the same genres across stations. Micro level analyses will focus on fluctuations between individual adjacent programme

elements. Each discreet programme segment is logged in duration, integrated loudness, loudness range, genre, and time of day, with additional subjective assessment of the audio captured as notes (noticeable distortion, panning bias, etc.). This data allows for the following analysis:

- All segments combined and weighted for duration gives the overall station integrated loudness and LRA values. These values combine to provide the probability of inter-station loudness fluctuations when switching between stations.
- Using similar methods, single integrated loudness values are calculated for the four main programme types: talking, advertisements, other interstitials (called ‘links’ in the results) and music. The purpose of this is to search for any trends in the balance between these elements, as well as testing the common laypersons’ belief of ‘advertisements are louder’.
- By comparing the integrated loudness value of each discreet segment to that of the following segment, intra-station fluctuations can be quantified. These will be presented in a number of ways per station. A scatter-plot of all discreet adjacent integrated loudness values. A histogram of fluctuation data, also indicating data bins falling outside the loudness ‘comfort zone’, in regions where 50% of listeners, and 95% of listeners would take corrective action. The overall proportion of fluctuation within and outside the ‘comfort zone’ will be presented as a stacked bar chart per station.
- Data from the ‘zap tests’ will be analysed and presented in the following ways. Actual simulated inter-station fluctuations will be plotted as a composite of all individual zap-tests, including an aggregate of results. The discreet fluctuations will also be plotted according to proportion outside and within the ‘comfort zone’ parameters.
- Lastly, discreet LRA value fluctuations will be analysed according to listener dynamic range tolerance parameters, and presented as a distribution. This adds value as two stations may have identical station LRA values, but the station with a narrow distribution of segment LRA values will indicate more dynamic range processing on transmission level.

All data capture logs are imported into various spread sheets in Google Sheets⁴, both for backup purposes, and for the powerful data analysis and chart generation features. Further data analysis methods will be discussed in the next chapter alongside the research findings.

⁴ The Google Sheets application is a free, online spread sheet that is very similar to Microsoft Excel.

Chapter 4

High level research findings

This chapter presents the findings of loudness measurements textually and graphically, with a focus on the long-term integrated loudness values, presented as separate sections for terrestrial television and radio broadcasts. Within each of these sections, the results will be presented as integrated loudness values and other metrics for the entire station or channel. In this study, more than 2000 individual audio segments were analysed, however this chapter will start with high level results only. It will only present loudness data as measured and aggregated, and in that way is different from Chapter 5 where the fluctuation data is extrapolated and evaluated according to the listener tolerance metrics, including ‘zap test’ methodology and scenarios. This chapter focuses macro level dynamics and description, and on inter-station fluctuations, setting out to answer the main research question, “*does South African free-to-air broadcast audio contain loudness fluctuations of a magnitude that exceeds previously determined limits of listening comfort?*”

4.1 Radio findings

Comprehensive aggregated loudness data captured for radio stations is given in Table 5 below:

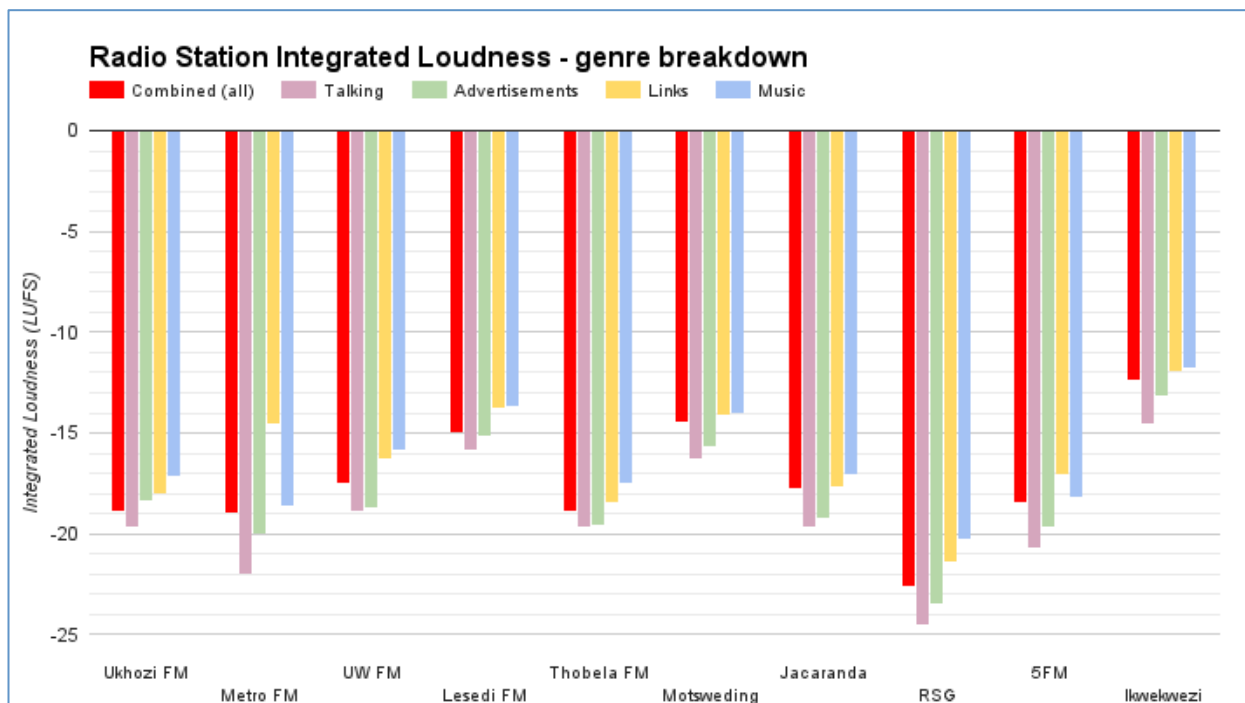
Table 5: Station integrated loudness and loudness range values for top 10 radio stations

STATION	STATION OVERALL				TALKING		ADVERTISEMENTS		LINKS		MUSIC	
	Rank	Loudness (LUFS)	LRA (LU)	Segments	Loudness	LRA	Loudness	LRA	Loudness	LRA	Loudness	LRA
UKHOZI	1	-18.9	4.6	111	-19.7	3.2	-18.4	3.1	-18	8.1	-17.2	3.4
METRO	2	-19	10.8	118	-22	11.9	-20	9.9	-14.6	11.5	-18.6	6.1
UW FM	3	-17.5	5.8	95	-18.9	4.6	-18.7	3.2	-16.3	6.7	-15.9	2.6
LESEDI	4	-15	5	123	-15.9	4.1	-15.2	3.4	-13.8	8	-13.7	3.8
THOBELA	5	-18.9	4.9	121	-19.7	4.5	-19.6	3.5	-18.5	7.7	-17.5	2.7
MOTSWEDING	6	-14.5	4	107	-16.3	4.4	-15.7	2.6	-14.1	7.7	-14	2.4
JACARANDA	7	-17.8	5.1	118	-19.7	4.9	-19.2	4.3	-17.7	8.5	-17.1	3.1
RSG	8	-22.6	9.6	104	-24.5	6.9	-23.5	9.6	-21.4	13.8	-20.3	8.6
5FM	9	-18.5	8.7	101	-20.7	9.3	-19.7	9.9	-17.1	9.8	-18.2	7.2
IKWEKWEZI	10	-12.4	9.1	91	-14.6	9.4	-13.2	15.1	-12	9.9	-11.8	2.2
RADIO 2000*	/	-21.5	7.2	49	-21.4		-25.4		-18		-21.9	
POWER FM*	/	-26.1	3.4	28	-26.3		-24.5		-24.4		/	
CLASSIC FM*	/	-18.5	5.8	37	-20.2		-18.4		-19.6		-18.3	
SAFM*	/	-20.5	7.5	31	-20.6		-19.8		-16.6		/	

* Less data was captured for these four stations. Enough to provide robust integrated loudness ratings, but not enough for genre-specific LRA ratings. Due to station programming, both Power FM and SAfm did not broadcast enough music segments for robust genre loudness ratings. Refer to section 3.2.3 for station selection details.

The number given for ‘segments’ refers to the total amount of discreet programmes logged; not data logging blocks. In loudness normalisation, R128 provides for a ± 1 LU tolerance in integrated loudness measurements (EBU 2011a: 4). The top 10 stations on Table 5 already reported station loudness values within this tolerance after the first 15-20 segments when compared to the final values. However, more extensive logging was done to enable better comparison between day of the week, time of day, and most importantly, between programme types or genres as given in Figure 4.1.

Figure 4.1: Station integrated loudness with genre breakdown for top 10 radio stations

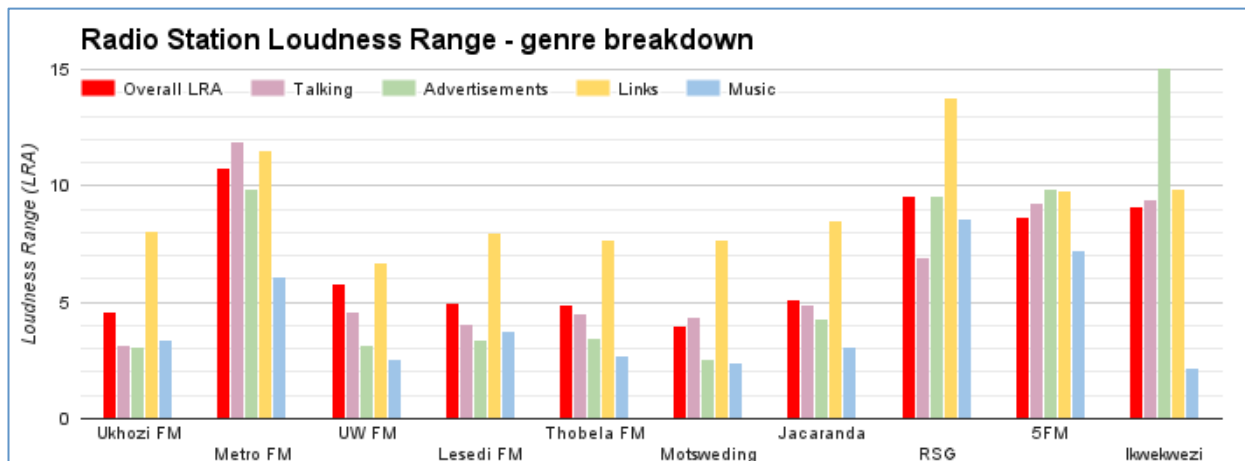


As evident from Table 5 and Figure 4.1, there are large differences between the station loudness of the top 10 stations, and even more so when additional stations are included. Most stations can be considered as ‘loud’, with Lesedi, Motswedding and Ikwelowezi being ‘very loud’. Ikwelowezi is by far the loudest station with -12.4LUFS overall loudness with RSG the softest at -22.6LUFS – also the only station within the R128 -23LUFS ± 1 LU tolerance. When including the extended stations, Power FM was the overall softest at -26.1LUFS. This gives overall ranges of 10.2LU and 13.7LU respectively, well outside the key metric 8LU ‘comfort zone’ of +3LU and -6LU (given in Table 2).

Note how *advertisements* are generally not the loudest programme elements, but rather consistently the second softest after *talking*. For most stations, the *music* elements are in fact the loudest, followed by *links*, with the exceptions of Metro FM and 5FM where the *links* are louder than the *music* elements.

For loudness range values, the metric provided by Lund (2006a; 2008; 2013) presented in Figure 2.8 can be summarised in the following loudness range tolerance values: in-flight entertainment 6LU, car 8LU, portable player 10LU, bedroom or kitchen 12LU, living room 20LU, and home theatre system 28LU. Figure 4.2 provides the overall loudness range of each of the top 10 stations and also gives a breakdown per genre.

Figure 4.2: Station loudness range with genre breakdown for top 10 radio stations



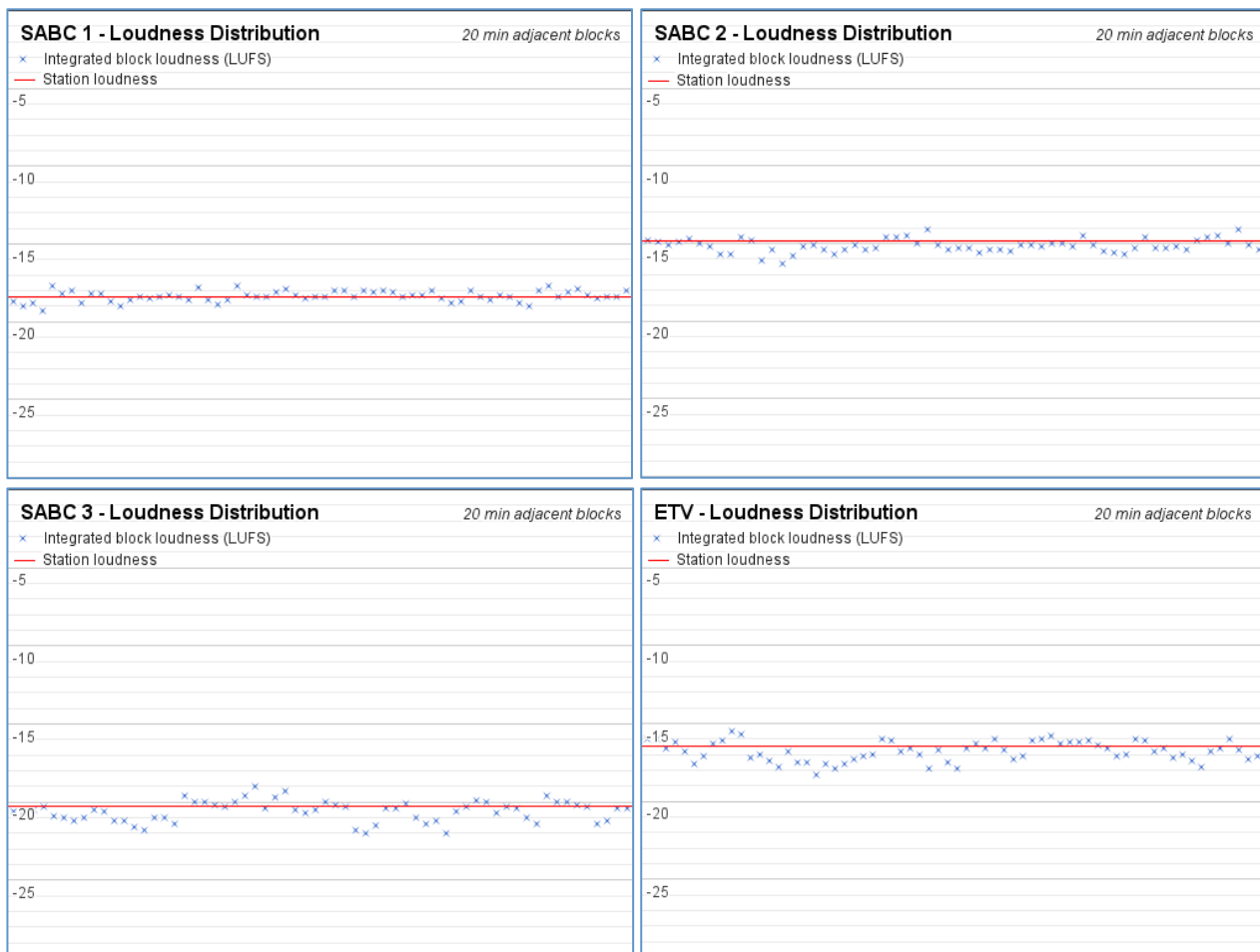
As evident in Figure 4.2, more than half of the top 10 stations have overall loudness range values that are extremely compressed, and seem to cater for the lowest common denominator of noisy in-transit listening. With the exception of Metro FM, all other top 10 stations have loudness range values appropriate for listening in the car.

One element that stands out is that, in most stations, the loudness ranges of *link* elements are considerably larger than other content types. The author contends that this is a result of a combination of *links* generally being louder, but also short in length. As explained in section 2.3.2, loudness range is a measure of macro-dynamics, and is less useful for shorter segments. Also note that in most stations, the loudness range for music segments is the smallest, with stations like Motsweding and Ikwewezi having extremely compressed music with values close to 2LU. Further analysis of these elements in Chapter 5.

4.2 Television findings

It should be noted that less results are presented in the television section compared to radio. This is not only due to the small amount of public stations compared to radio, but also since early investigations revealed that analogue television broadcasts employ gain control systems (see Figure 4.4) and aggressive audio limiting that, for the most part, also prevent problematic loudness fluctuations on the viewers' side, regardless of fluctuations in the original content (refer to Figure 4.3). The side effects of these are discussed at the end of this section.

Figure 4.3: Station loudness distributions for TV stations



Despite the lack of problematic fluctuations between content segments in a station, the overall station loudness values are sufficiently far apart, that fluctuations outside the 'comfort zone' will occur when switching between stations. The probability of these fluctuations will be high since all stations have a very narrow overall loudness range (refer to Table 6). SABC 3 was measured to be the softest at -20.2LUFS with SABC 2 the overall loudest at -14.8LUFS, giving a range of 5.4LU between the extremes. Even though not a single measured loudness fluctuation in the television data fell outside the viewer 'comfort zone', fluctuations do still

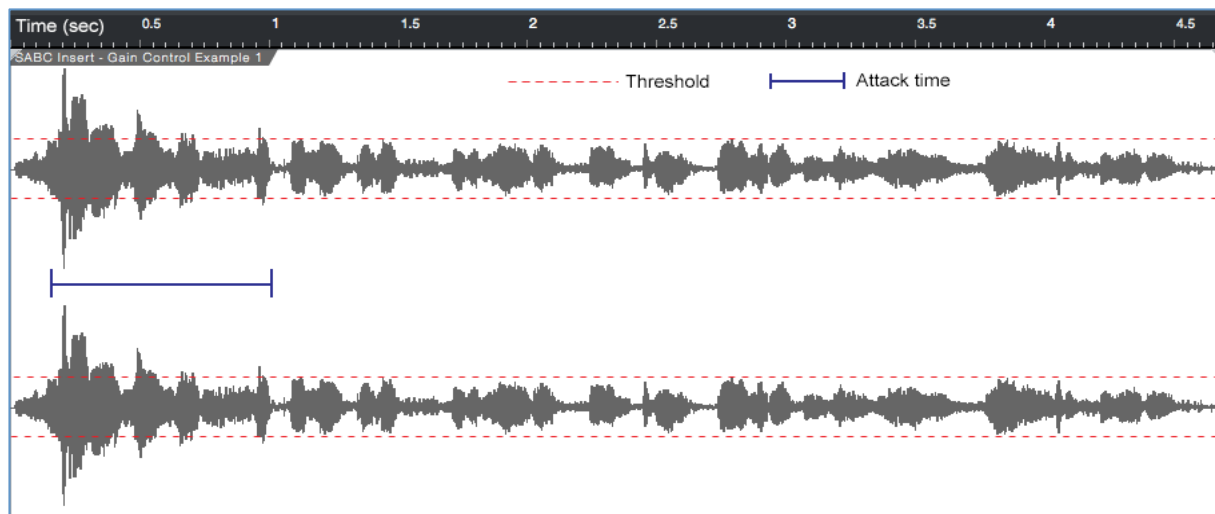
occur. Referring to both Table 6 and Figure 4.3, it is clear that SABC 1 has by far the smallest fluctuations corresponding to the smallest loudness range, with SABC 3 and e.tv containing fluctuations of a higher magnitude.

Table 6: Station loudness and loudness range for TV stations

	Loudness (LUFS)	LRA (LU)
SABC 1	-18.9	4.6
SABC 2	-14.8	4.9
SABC 3	-20.2	5.5
e.tv	-16.3	6.8

Note that integrated loudness values for TV stations cannot directly be compared to radio station values due to different signal paths. TV signals went through an additional gain stage, and while careful calibration ensured robust relative levels, they cannot be treated as absolute full-scale levels. Loudness range values on the other hand can be compared across the board.

Figure 4.4: Waveform example of automatic gain control in television broadcast



The content of the waveform shown in Figure 4.4 above, was from a programme link with a uniform voice-over with consistent background music. It is clearly evident from playback and the waveform that gain control is being applied, and it seems as the attack time is somewhere between 0.5 and 1 second. This would be the most likely explanation for the lack of intra-station fluctuations in television broadcast audio, but also leads to very narrow dynamic range, and lack of transparency in the broadcast chain. Content producers do not have the final control over the aesthetics of their mixes. While these fluctuations might still cause viewer annoyance, they are unlikely to result in viewer action as the level does reduce reasonably quickly. The annoyance factor of such short instances is beyond the scope of this study.

Chapter 5

Detailed findings, analysis and discussion

This chapter answers the two subsidiary research questions, evaluating the nature and extent of loudness fluctuations described in the previous chapter, and the platform appropriateness in terms of loudness range. This is achieved by placing the integrated loudness results of individual adjacent fluctuations within the key metric of listener ‘comfort zone’. While the previous chapter has already confirmed the likelihood of encountering inter-channel loudness fluctuations based on integrated station loudness, this section will start with an overview of intra-station fluctuations across all stations, followed by separate discussions on each station. The results from the simulated ‘zap tests’ are provided and discussed. Thereafter it is explored whether ‘zap test’ methodology could be efficient for loudness monitoring, and whether LRA values have any correlation with intra-station fluctuation proportions. This chapter has a greater focus on radio broadcasts.

5.1 Intra-station fluctuations in radio loudness

Loudness fluctuations outside the listener ‘comfort zone’ were found at each of the top 10 radio stations under discussion, but the proportion and extent varied per station. Table 7 below gives an overall view of these fluctuations, and refers to the difference in integrated loudness of adjacent programme blocks (or loudness fluctuations between segments). It combines 50% and 95% action sections as both of these are considered outside the ‘comfort zone’, as explained in section 2.3.1.

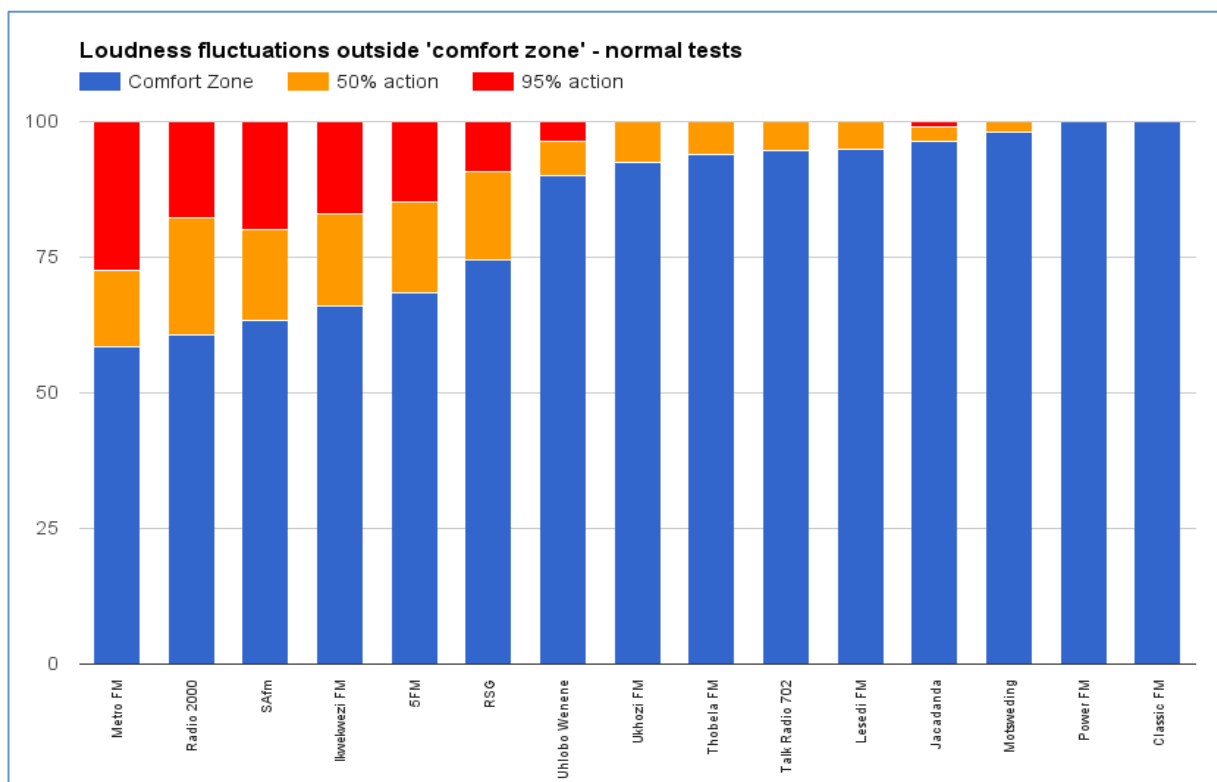
Table 7: Percentage of fluctuations outside ‘comfort zone’ – top 10 radio stations

	Fluctuations outside ‘comfort zone’
Metro FM	42%
Radio 2000	39%
SAfm	37%*
Ikwewezi FM	34%
5FM	32%
RSG	26%
UW FM	10%
Ukhozi FM	8%
Thobela FM	6%
Talk Radio 702	5%
Lesedi FM	5%
Jacaranda	4%
Motsweding	2%
Power FM	0%*
Classic FM	0%*
<i>*smaller dataset</i>	

The results are based on discreet programme blocks that varied in length. A full song, or discussion programme is considered a single programme block, as a listener would choose an appropriate listening level based on the content of such a coherent block. Fluctuations occur in one of two ways. Firstly, when the integrated loudness of the next, adjacent block is considerably louder or softer, listeners would take action to compensate for this. Secondly, when the content within a single programme block varies too much, this would also cause annoyance, but it is more difficult to counteract. This section explores the former, while the section on LRA explores the latter.

As presented in Table 2 in section 2.3.1, it was previously determined that 50% of listeners would take action if adjacent material fluctuates to be 3LU louder or 6LU softer, while 95% of listeners would take action for loudness variations of 5LU louder and 8LU softer, respectively. Other, smaller fluctuations not only occur, but are noticeable, but these all fall within the so-called ‘comfort zone’. Figure 5.1 below provides a further breakdown, per station, sorted by most fluctuations, showing the proportion of fluctuations falling in each of these categories. Notice how Metro FM has an extraordinary 42% combined percentage of problematic fluctuations, nevertheless, even a small percentage in any of the fluctuations is still considered to be problematic.

Figure 5.1: Proportion of fluctuations outside listener ‘comfort zone’



Note that all of the top 10 stations have some intra-station fluctuations outside the ‘comfort zone’. The two stations where 0% is given, Power FM and Classic FM, were both part of the additional 4 stations included in the comprehensive part of the study designed to analyse additional elements, and both had much smaller datasets. However, both these stations apply a large amount of dynamic compression, which is a likely reason for the lack of intra-station fluctuations. This will be discussed in greater detail in the next section on individual stations.

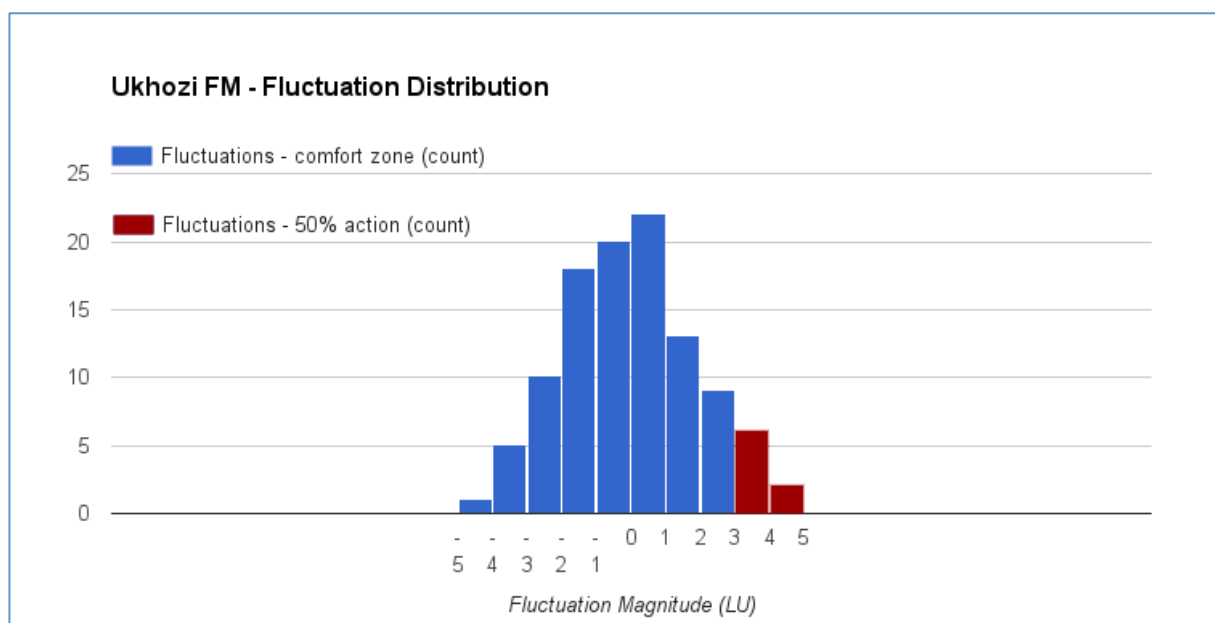
5.2 Discussion of individual radio stations

This section will provide a detailed view of intra-station fluctuations by discussing each radio station separately. It provides fluctuation distributions for each station and also includes other observations made during data capture. Please refer to Appendix C for full graphs plotting the loudness distribution of each adjacent programme element for each station. All distribution graphs have been presented according to the same scale for ease of visual comparison.

5.2.1 Ukhozi FM

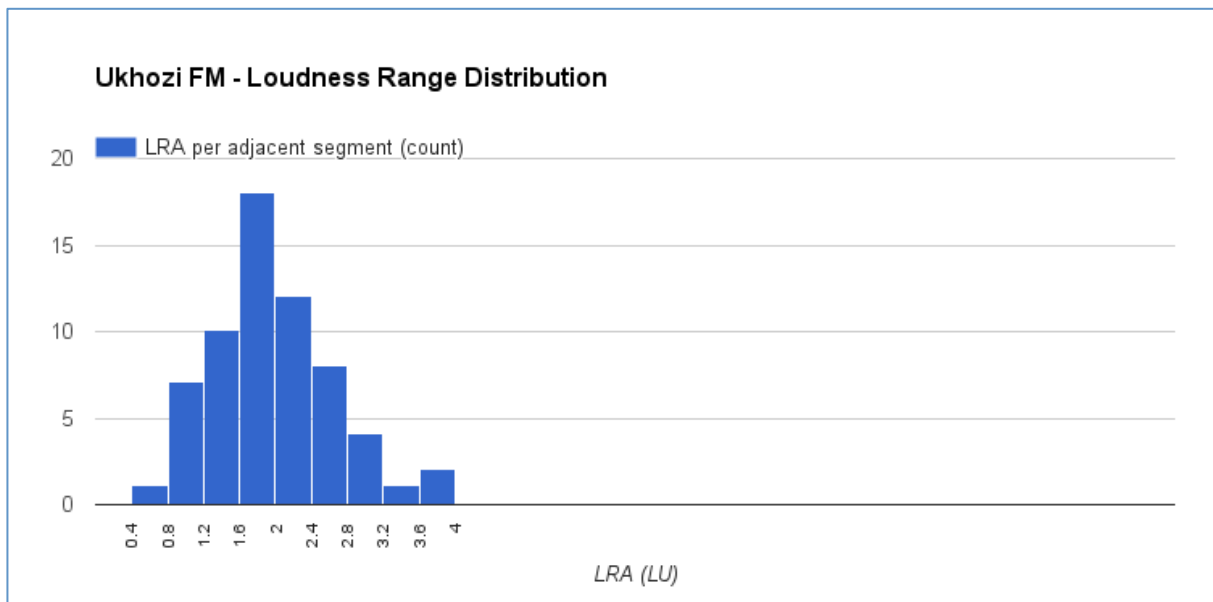
Ukhozi FM is the station with the highest national audience figures, consistently scoring 20% and above in the RAMS results. The integrated station loudness was measured at -18.9LUFS and the station LRA at 4.6LU. Of the 111 individual adjacent segments, 8% fell outside the listening ‘comfort zone’, all of which within the +3-5LU range of 50% listener action as indicated on Figure 5.3. Content on the station was generally evenly levelled but highly compressed.

Figure 5.3: Ukhozi FM Fluctuation Distribution



A number of operator errors were observed during data logging, most frequently music tracks and theme songs being cut abruptly short. In some instances even advertisements were cut short. Some of the weekday talking segments presented gross clipping distortion when the presenters suddenly spoke loudly or laughed. Similar to the overall station loudness range, the distribution of the LRA values of individual segments have a very narrow distribution with the majority of segments having an LRA of around 2LU and less, as seen in Figure 5.4. This is a good indicator of the high amount of dynamic compression applied before transmission, and would explain why all segments and the overall LRA is narrow enough to be platform appropriate even for the noisiest environment.

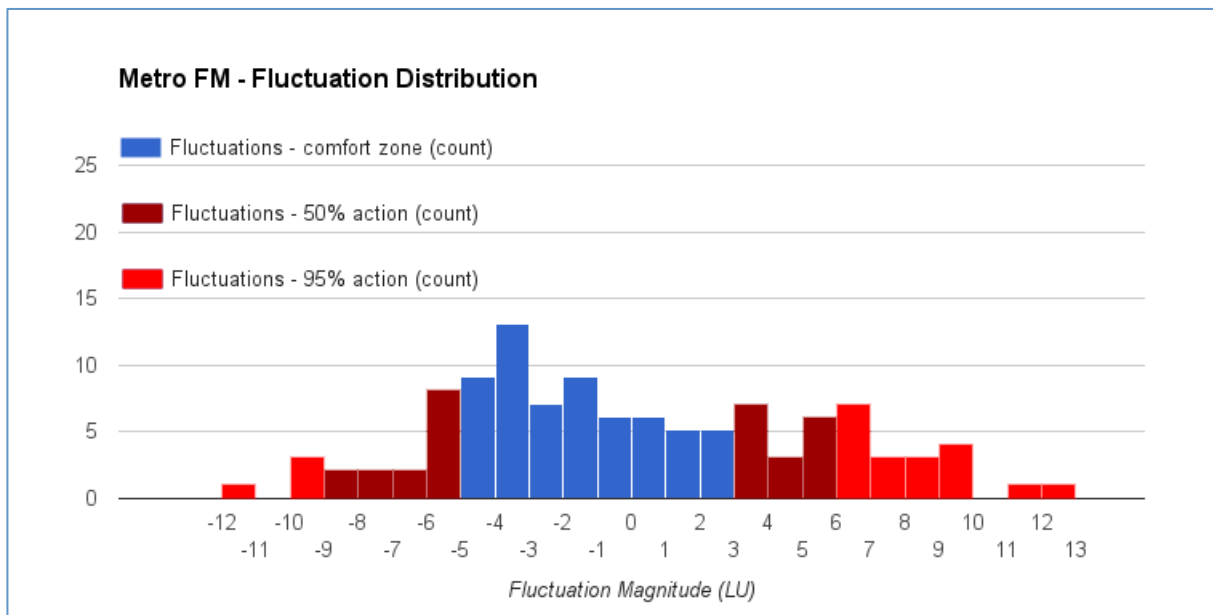
Figure 5.4: Ukhozi FM Loudness Range Distribution



5.2.2 Metro FM

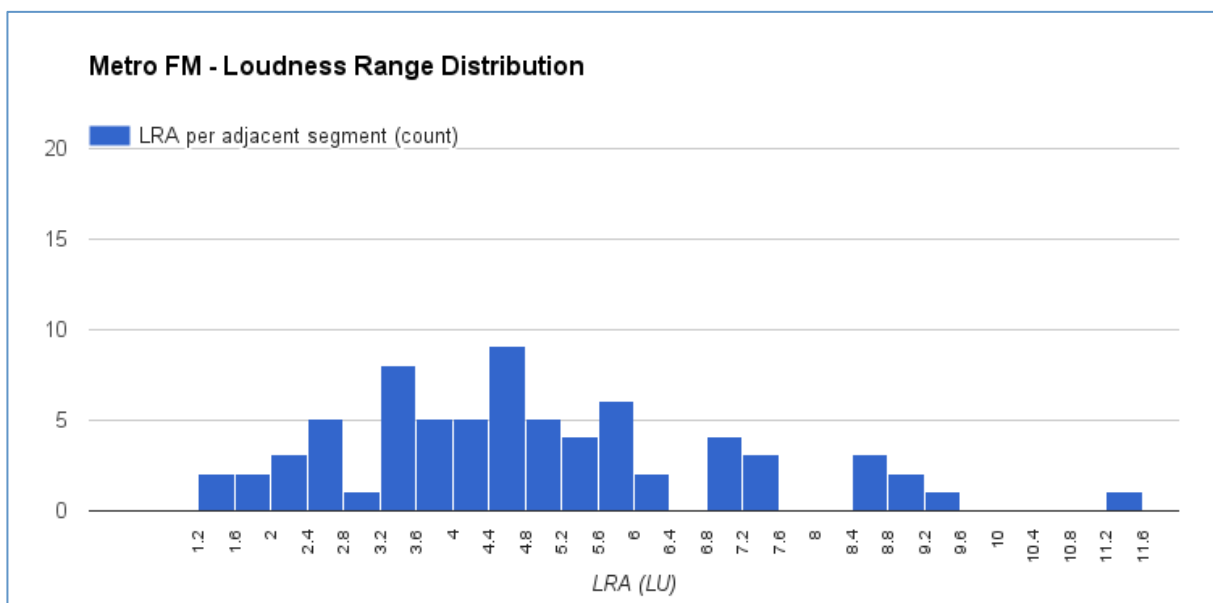
Metro FM has a very similar station loudness value at -19LUFS, but all other elements are of stark contrast. Of all the station observed and measured, it had by far the most and greatest fluctuations across all the metrics. A staggering 42% of adjacent programmes exhibited loudness fluctuations outside the listening ‘comfort zone’, just over 27% of which fall in the 95% listener action range as presented in Figure 5.5. The biggest element causing offending fluctuations is Metro FM’s own interstitials. These *links* were almost always transmitted considerably louder than adjacent content, or even on top of already loud music. In two measurements, these fluctuations were in excess of +10LU.

Figure 5.5: Metro FM Fluctuation Distribution



In addition to the large amount of adjacent segment fluctuations, the loudness range also varied considerably as presented in Figure 5.6. An exceptional amount of operator errors were observed during data capture. These included frequent abrupt stops of all programming types, advertisements being played back while the presenter is still talking, stopped again, and then restarted, instances where the background music was played louder than the talking, resulting in near impossible speech intelligibility, and frequent and very loud popping plosives. Microphones were often kept live underneath other content (with a presenter in one instance heard taking a personal phone call while an advertisements was aired).

Figure 5.6: Metro FM Loudness Range Distribution



5.2.3 UW FM

UW FM has an integrated station loudness of -17.5LUFS, and station LRA of 5.8LU. Of all adjacent programmes, 10% exhibited loudness fluctuations outside the ‘comfort zone’. The single outlier fluctuation on Figure 5.7 was caused by an operator error, playing a segment at a very low level, and then playing the same segment again much louder. Overall advertisement levels were very closely matched to those of talking segments.

Figure 5.7: UW FM Fluctuation Distribution

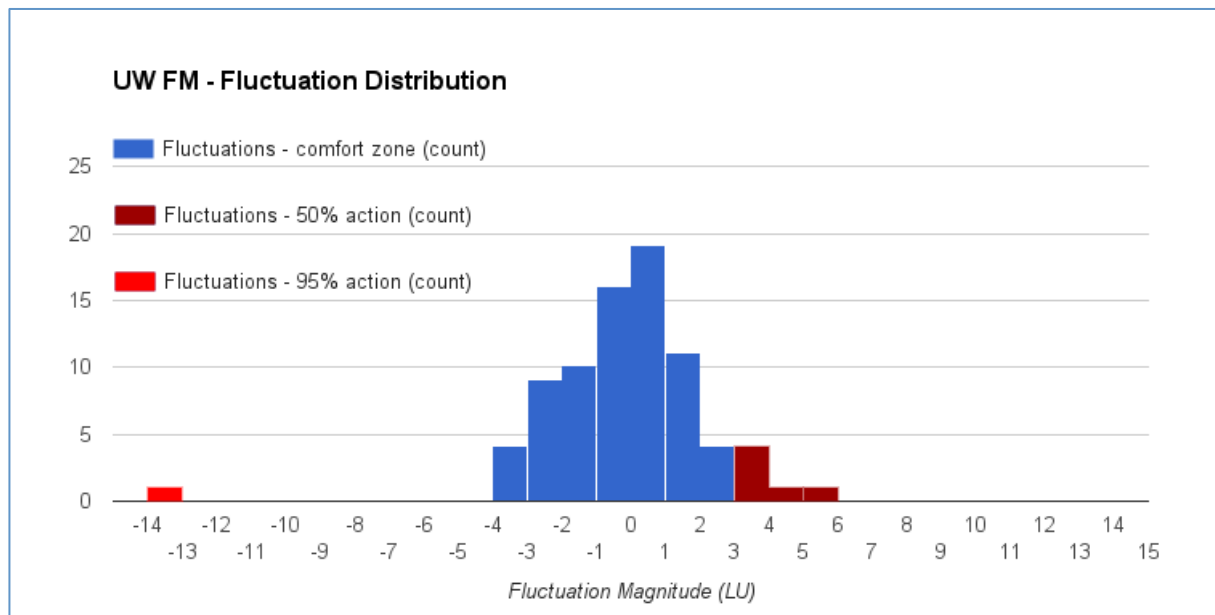
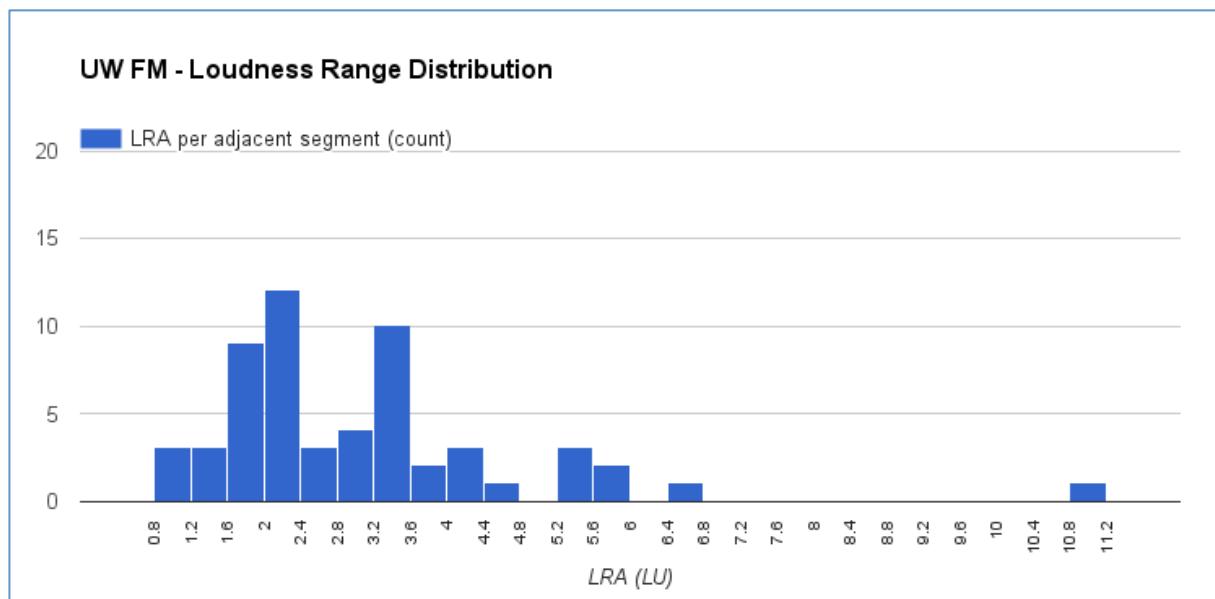


Figure 5.8: UW FM Loudness Range Distribution



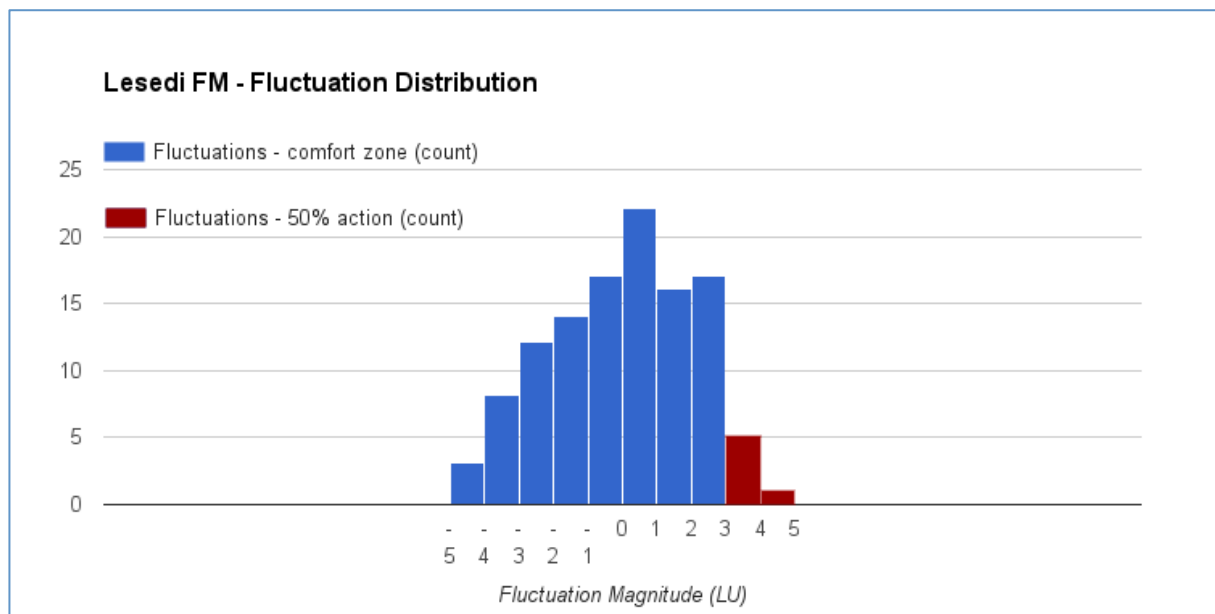
As evident in Figure 5.8, the loudness range of segments were usually quite narrow, but with a number of higher range outliers. The music aired on this station had an extremely narrow overall

LRA of just 2.6LU. This could partly be due to the genre of music being aired, but additional dynamics processing on the transmission was also audible. Of all the stations observed, UW FM exhibited the most problematic operations. This included extremely long gaps of silence between segments, and false starts of music playback with restarts. Buttons being pressed and talking in the background in the studio while other segments were playing was heard on a number of occasions. On the technical side, there was a consistent high frequency noise being transmitted whenever the studio microphones were live. This was not a general transmission problem as this noise was only present on in-studio talking segments. Additionally, gross distortion was frequently present on studio voices. This leaves the impression of an unprofessional radio station, even though it is a commercial station with the 3rd highest listenership in the country on national level.

5.2.4 Lesedi FM

Lesedi FM measured to be the third loudest station of the top 10, with -15LUFS station loudness, and a station loudness range of 5LU, and only 5% of fluctuations falling outside the ‘comfort zone’. It was clearly audible that music playback was consistently louder than other programme types, with integrated loudness of -13.7LUFS for all music segments combined. Links were also consistently louder than talking, measuring -13.8LUFS.

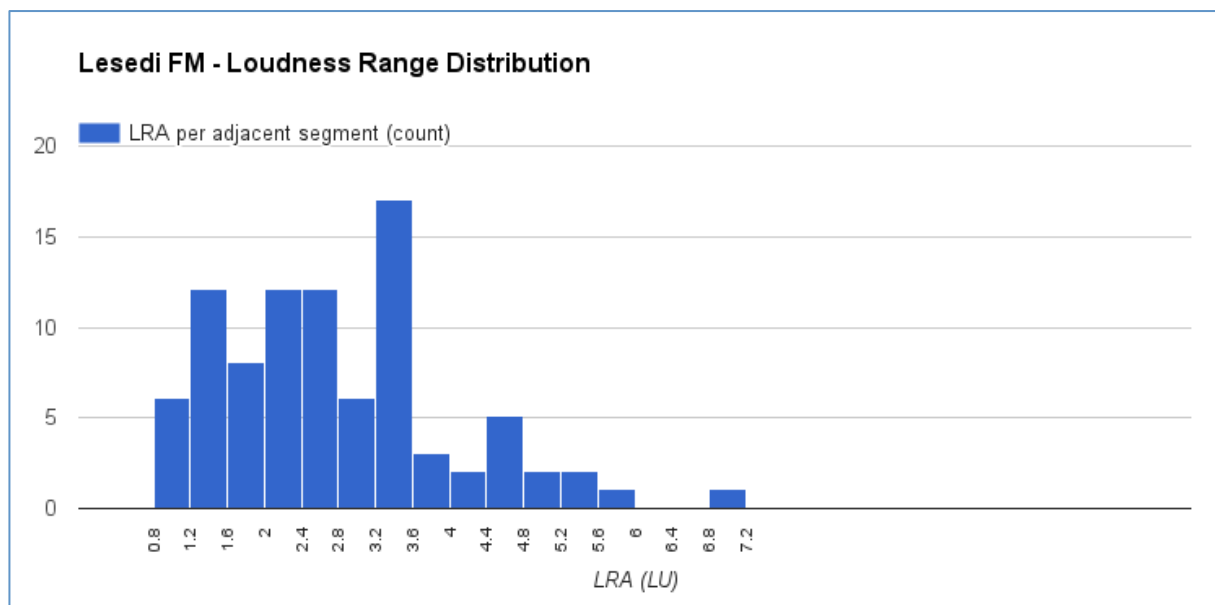
Figure 5.9: Lesedi FM Fluctuation Distribution



The distribution in Figure 5.9 corresponds with the narrow station LRA of 5LU, and confirms that all content is highly compressed before transmission – which was audible during data gathering.

The quality of station operations was average, but long silences (more than 10 seconds at a time) observed between programming. Some distortion present on louder talking passages, but the most distinguishing feature of this channel (in terms of technical problems) was the settings on the music ducker. In this case, a ducker is a side-chain compressor that lowers the playback level of the music when a voice is talking in the studio. These are usually programmed to have a relatively fast attack time, and slower release so that the music gets lowered quickly enough when the voice starts, but does not return during pauses between sentences. The Lesedi ducker was set up with extremely fast attack and release times and very high gain reduction. This caused extremely wide ‘pumping’ (cyclical variation in loudness) in the music. Loud music was often played in the background while a presenter was talking, but every time the presenter hesitates ever so slightly, or even just take a breath, the music would pop back instantly and much louder than the voice. As shown in Figure 5.10, the loudness range was generally small, but with a number of segments exhibiting a higher LRA, explaining the station LRA of 5LU even though most segments had LRA values well under 4LU.

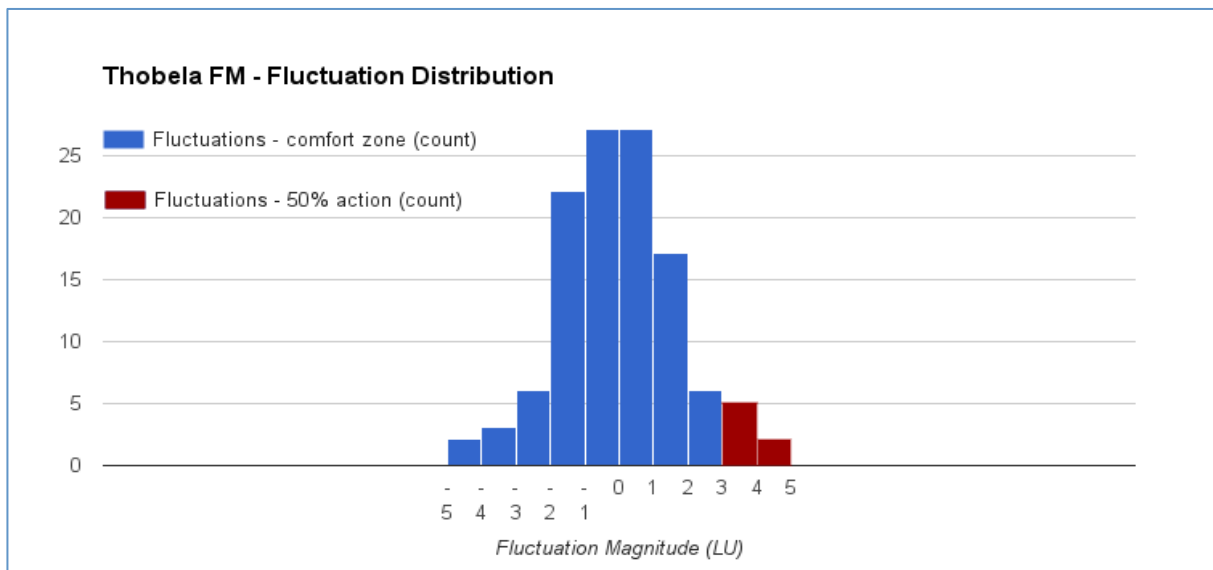
Figure 5.10: Lesedi FM Loudness Range Distribution



5.2.5 Thobela FM

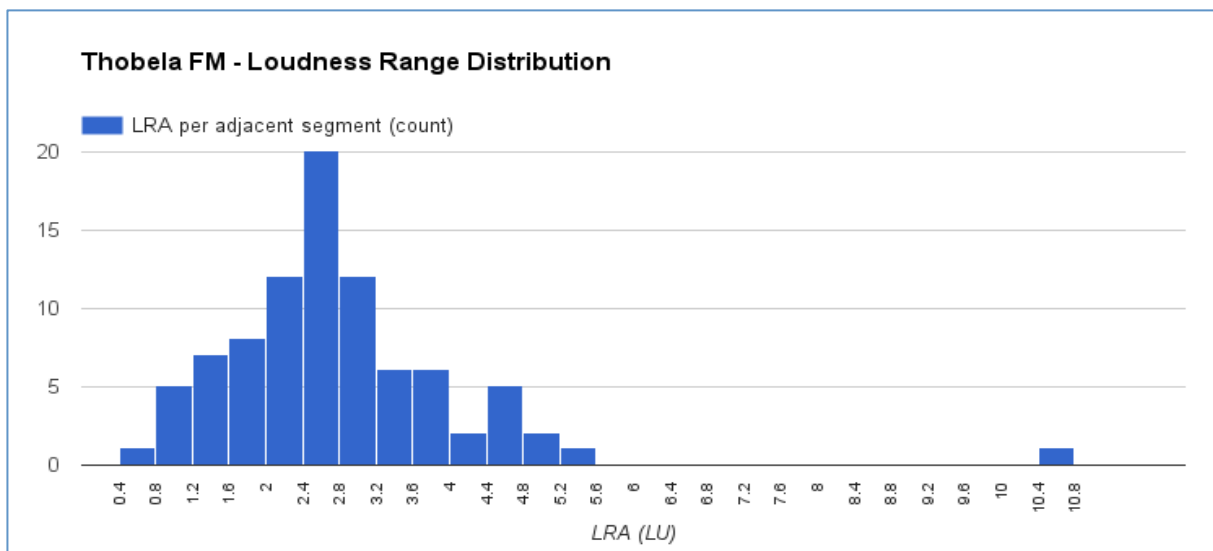
Thobela FM showed very similar characteristics to Ukhozi FM all round. Station loudness measured at -18.9LUFS with a station LRA value of 4.9LU. It has a heavily compressed transmission in general, with only 6% of fluctuations falling outside the ‘comfort zone’ as given in Figure 5.11. Overall advertisements levels were very closely matched to those of talking segments, with link segments somewhat louder.

Figure 5.11: Thobela FM Fluctuation Distribution



Music was consistently transmitted louder than other programme types, with an integrated loudness value of -17.5LUFS and an exceptionally narrow loudness range of just 2.7LU. This was also the most common LRA value across all content as shown in Figure 5.12. Thobela FM presented a high amount of controlling errors during data logging including the following: audible buzz on the transmission during all studio based content including a bias towards the left channel, gross distortion on audio links, long unintended silences, inserts and advertisements often cut short, and audible noise as buttons was pressed in the studio. This station also exhibited audio compression related pumping from an incorrectly configured ducker - similar, but not to the same extent as Lesedi FM. It did not sound like the technical aspects of this station were fully under control.

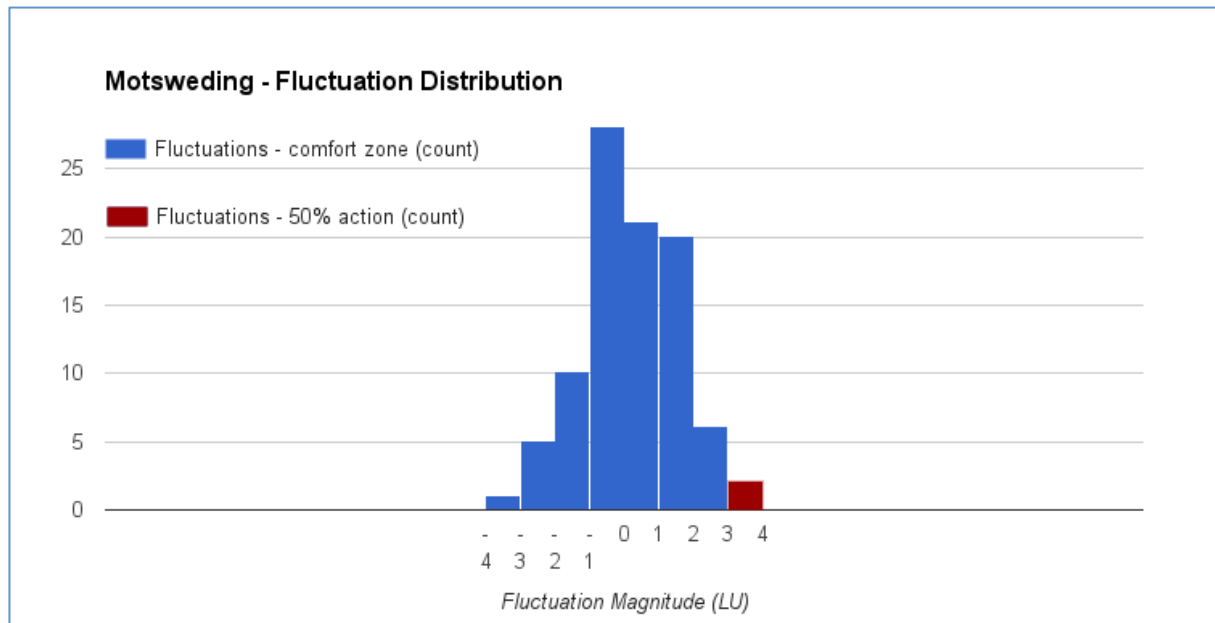
Figure 5.12: Thobela FM Loudness Range Distribution



5.2.6 Motswedding

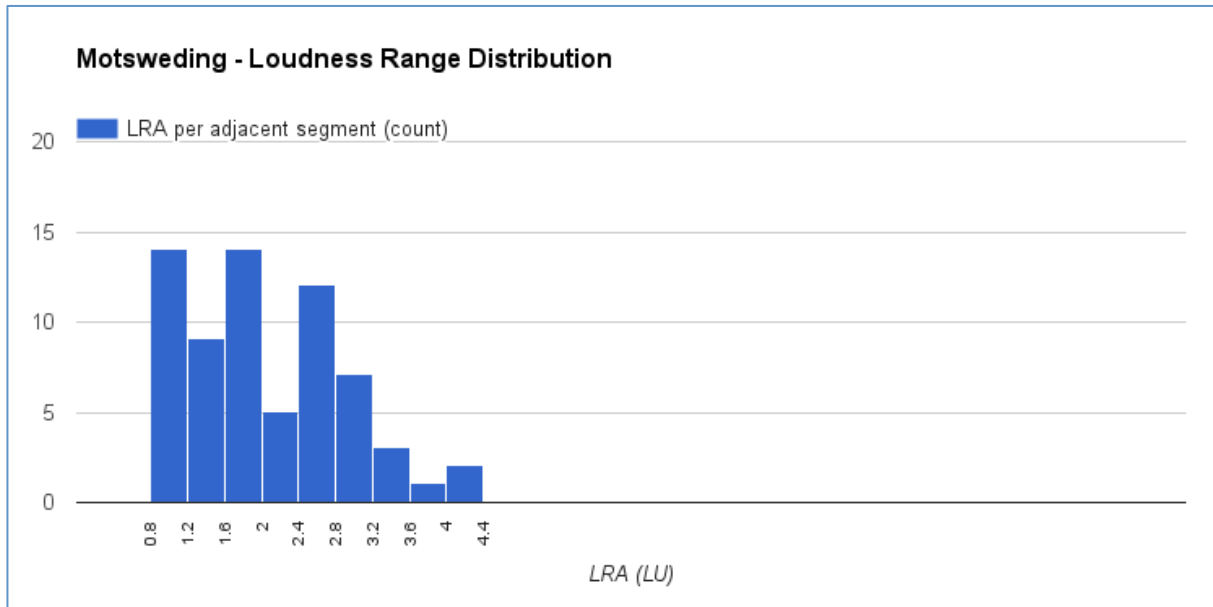
Motswedding was found to be the second loudest station with a station loudness value of -14.5LUFS, and also had the most compressed transmission overall, with a loudness range of just 4LU. Of all the stations with full datasets, Motswedding exhibited the least fluctuations outside the ‘comfort zone’ at a mere 2%, as depicted in Figure 5.13.

Figure 5.13: Motswedding Fluctuation Distribution



However, this station is a good example where a lack of fluctuations does not translate into a better overall listening experience. The low proportion of fluctuations is achieved by an exceptional amount of compression and peak limiting. Note on Figure 5.14 that the majority of programme segments had loudness range values less than 2LU. This station airs content from all programme types, and a significant amount of dynamics processing is required to achieve this. General technical control on this station was also exceptionally poor. Studio segments including a reasonably loud buzz, and talking segments frequently included plosive and signal pops, wind noise and awkward silences. The signal often exhibited gross distortion on all content types, but especially on talking segments in the studio. At one point, the background music and talking was processed with a series of delays that rendered the segment completely unintelligible. Motswedding also frequently included extreme music ducker pumping as explained in the section on Lesedi FM. Operations generally sounded unprofessional and not of a very high standard.

Figure 5.14: Motsweding Loudness Range Distribution



5.2.7 Jacaranda

Jacaranda had a station loudness value of -17.8LUFS and station LRA of 5.1LU. While the transmission is generally quite compressed, this station is an example of achieving a low proportion of fluctuations outside the ‘comfort zone’ (only 4%, as shown in Figure 5.15) while still generally maintaining the integrity of the audio. Unlike the previous few stations described, Jacaranda generally exhibited good technical operation. Talking and advertisements segments were generally well matched in loudness, and music and links, while louder, were also matched.

Figure 5.15: Jacaranda Fluctuation Distribution

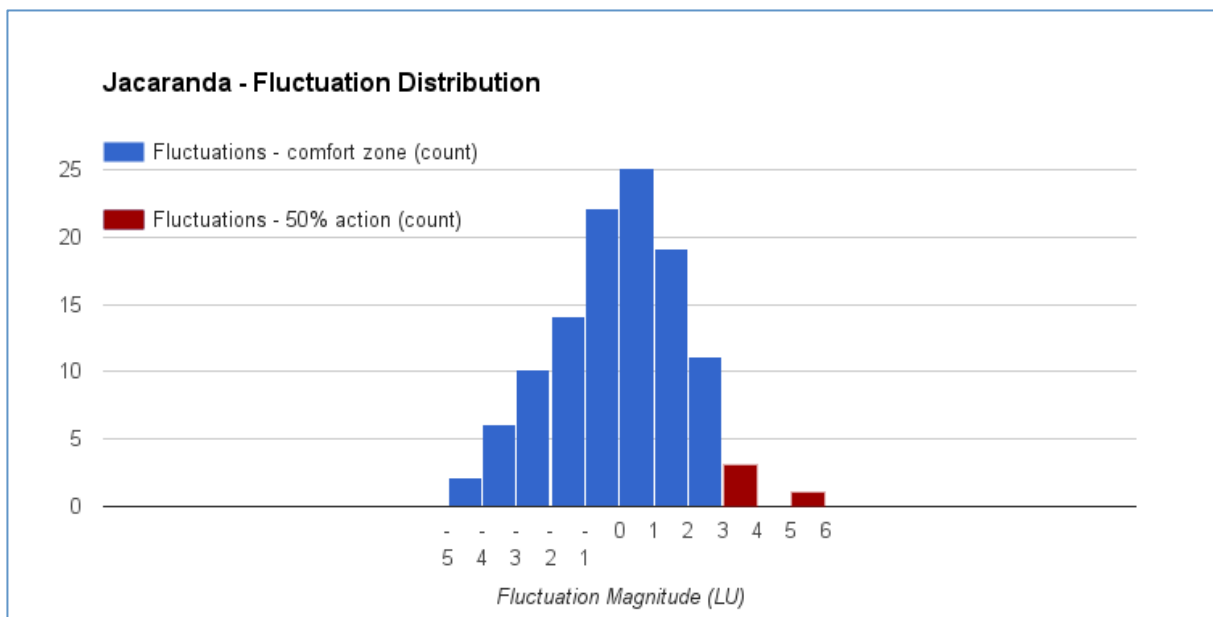
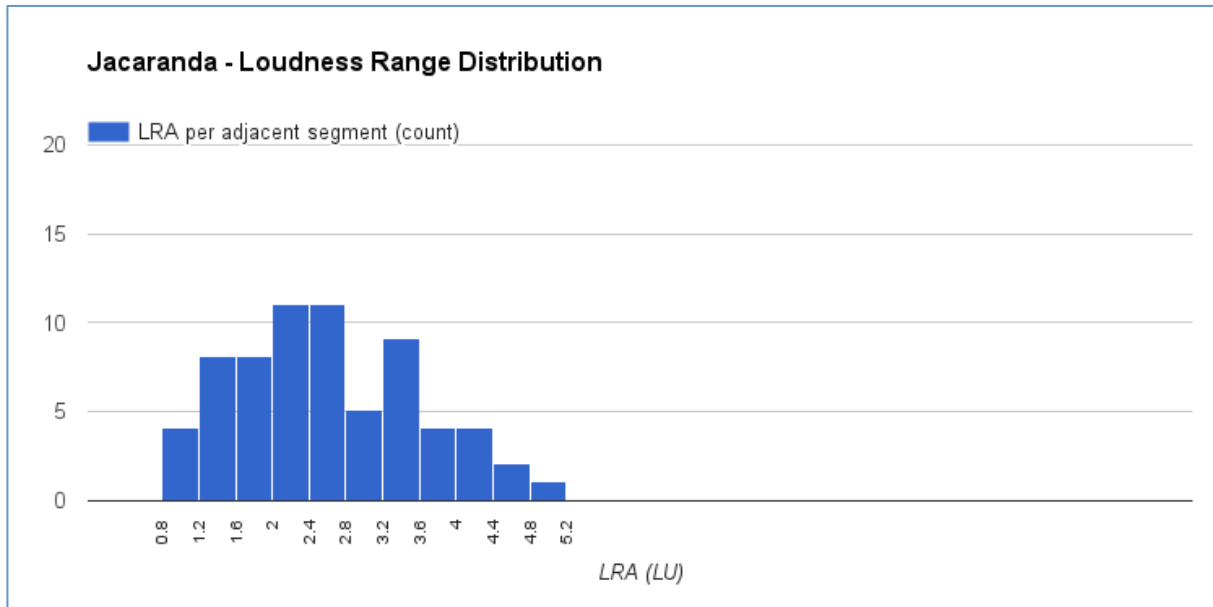


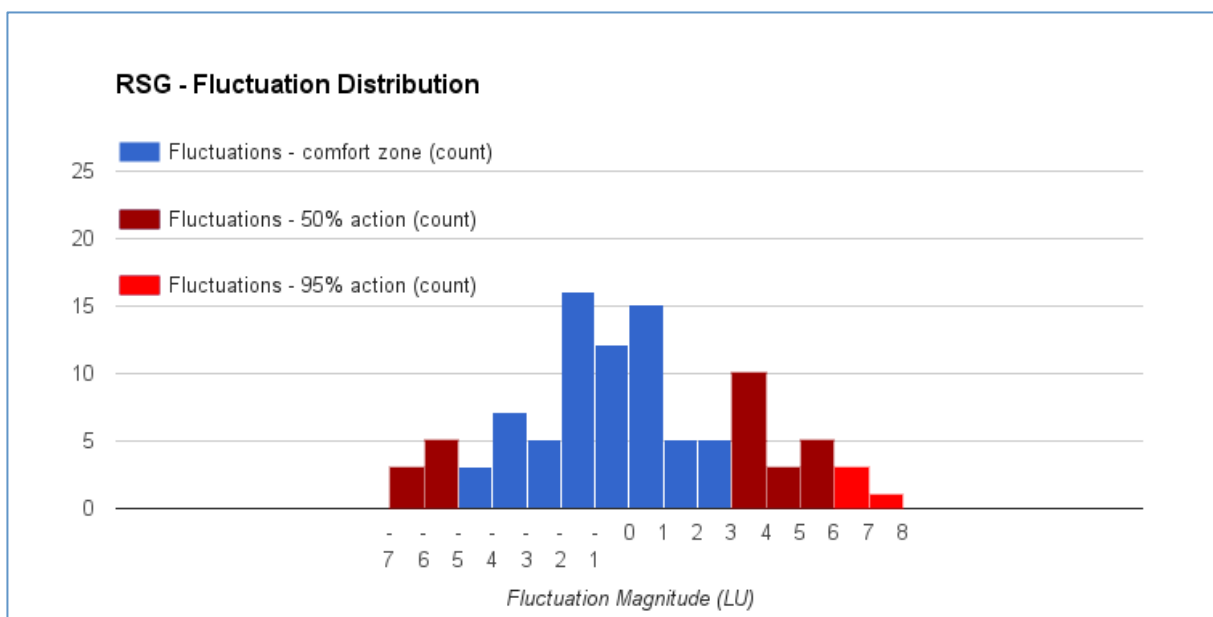
Figure 5.16: Jacaranda Loudness Range Distribution



5.2.8 RSG

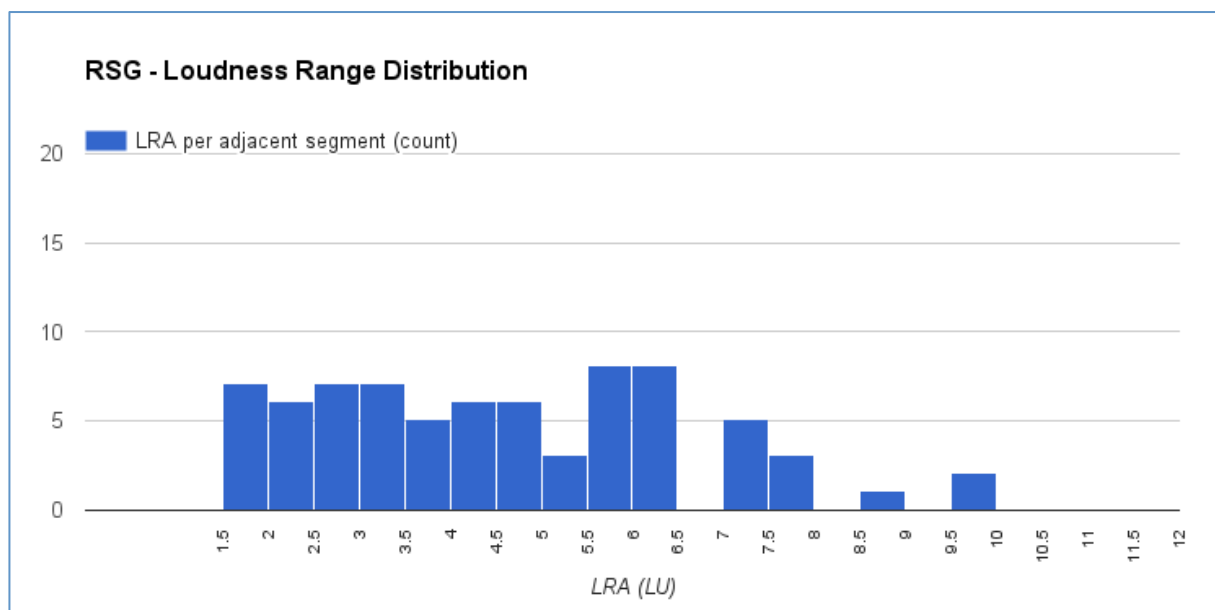
RSG was the softest station of the top 10 with station loudness of -22.6LUFS and station LRA of 9.6LU. While RSG broadcasts the closest to the EBU R128 recommended target level, this was probably due to the station leaving a more appropriate amount of headroom, and not due to the implementation of loudness normalisation. RSG exhibited 26% of fluctuations falling outside the ‘comfort zone’, including both positive and negative fluctuations, with nearly 10% of the fluctuations in the 95% listener action zone (refer to Figure 5.17).

Figure 5.17: RSG Fluctuation Distribution



Music segments (-20.3LUFS) were on average considerably louder than the average for talk segments (-24.5LUFS), which contributed greatly to the proportion of total fluctuations outside the ‘comfort zone’. The transmission was not generally as highly compressed as other stations with all content types measuring a loudness range of above 6LU (refer to Figure 5.18). The station generally had reasonably good technical control. As seen in Figure 5.18, no segments had loudness range values smaller than 1LU, and a reasonably large proportion had values higher than 6LU. RSG is the station in the top 10 that transmitted music segments with the widest average loudness range value of 8.6LU, and the station LRA is also the second widest of the top 10 stations after Metro FM. Many of the individual segments exhibited LRA values greater than would be appropriate for listening in a car, or even portable devices. These fluctuations are generally considered to be outside the loudness range tolerance for these specific listening environments.

Figure 5.18: RSG Loudness Range Distribution

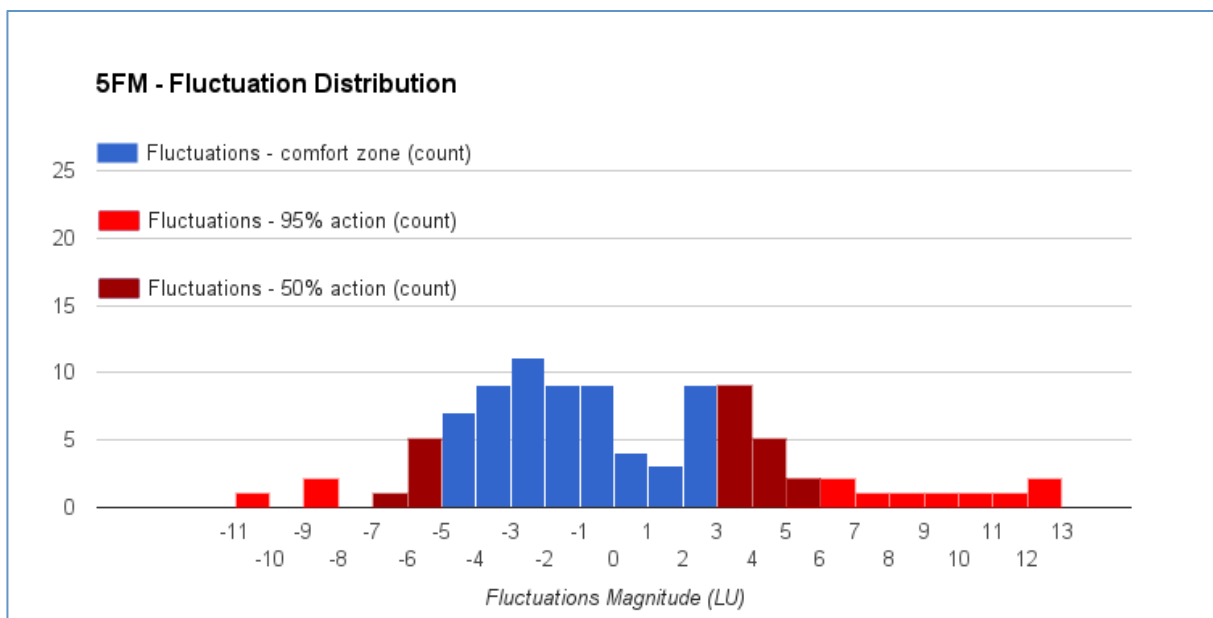


It would appear that RSG employs automatic gain control on in-studio talking segments, similar to the SABC TV stations, as illustrated in Figure 4.4. Talking segments after music segments would often start considerably louder, but would be attenuated within a second to a more appropriately matched level. This could be manual control, but the consistency of the attack time would suggest automatic adjustments.

5.2.9 5FM

5FM had an integrated station loudness of -18.5LUFS, with a station loudness range of 8.7LU. Nearly a third (32%) of all fluctuations fell outside the listening ‘comfort zone’, nearly half of which being in the 95% listener action zone as presented in Figure 5.19. Other than Metro FM, 5FM was the only station where music was not the loudest type of content, but rather links. A main reason for this is that they play out many links above music that is already playing back, without dropping the level of the music at all. For these links to be audible on top of the music, they are transmitted very loudly (with an average of -17.1LUFS integrated across all links).

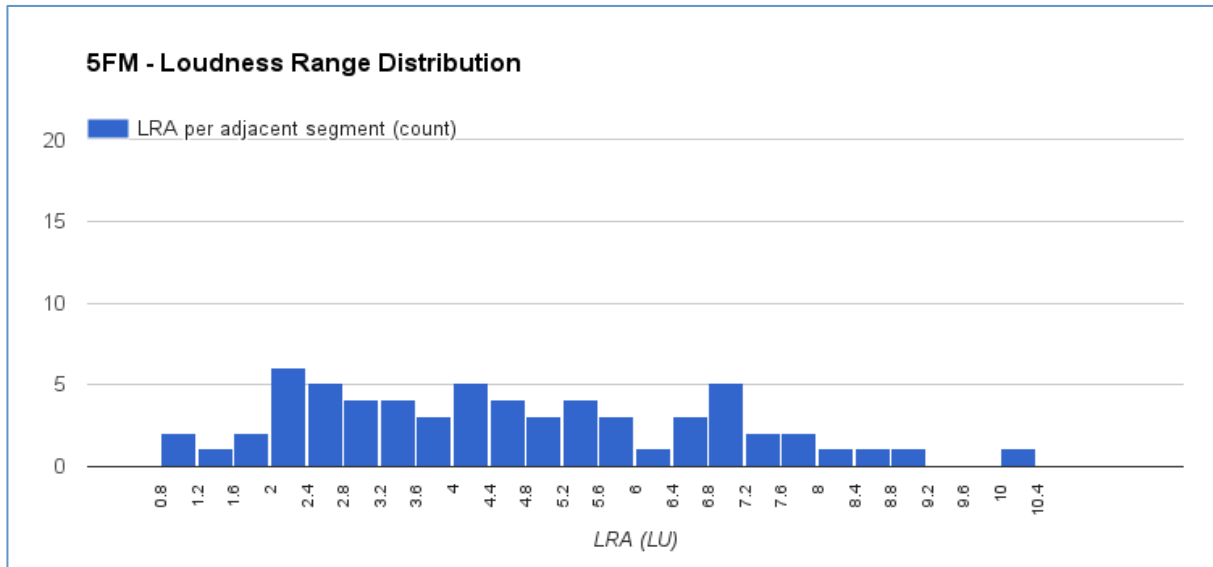
Figure 5.19: 5FM Fluctuation Distribution



The loudness ranges of segments are consistently quite large with talking at 9.3LU, advertisements at 9.9LU, links at 9.8LU and music at 7.2LU. There was also a wide distribution of segment loudness ranges, between 0.8LU and 10.4LU, as presented in Figure 5.20. Many of these segments would have ranges too wide to be appropriate for listening in transit, or on portable devices.

While 5FM has a large proportion of problematic fluctuations, both in integrated loudness and loudness range, the technical control of the station is generally good without any major operator errors observed during data logging.

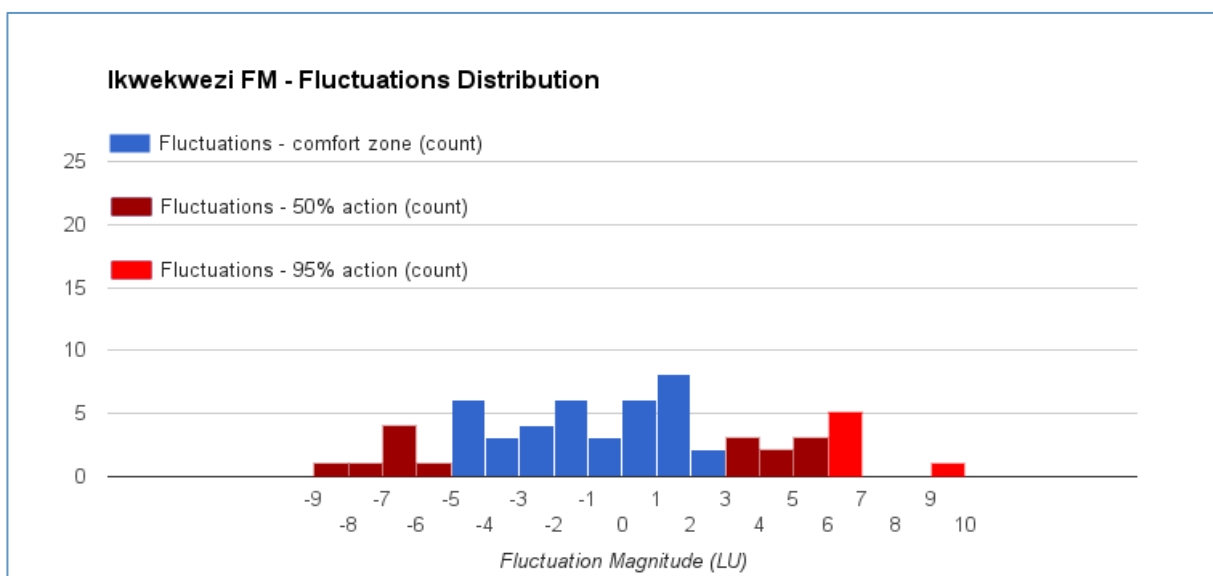
Figure 5.20: 5FM Loudness Range Distribution



5.2.10 Ikwekwezi FM

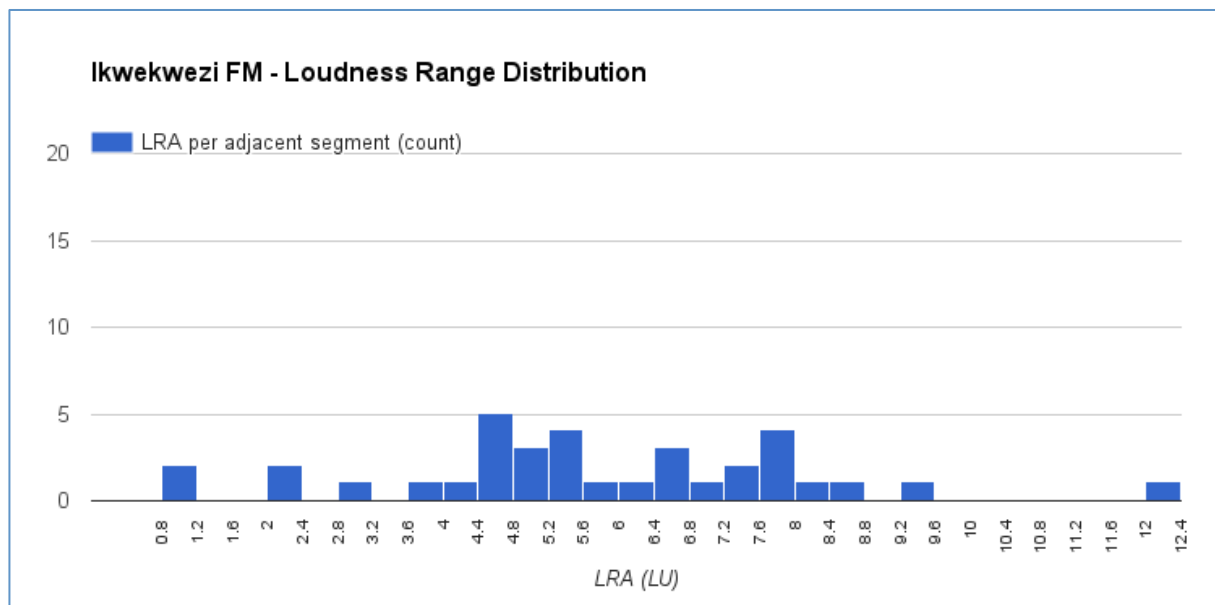
Ikwekwezi FM is an extremely loud radio station. With a station loudness value of -12.4LUFS and station LRA at 9.1, this station was found to be considerably louder than any of the other stations in this study. It requires a significant amount of compression and peak limiting to get the integrated loudness of a station to be this high. Despite the very large amount of dynamics processing, Ikwekwezi FM still exhibited a very high proportion (34%) of loudness fluctuations outside the listener ‘comfort zone’, approximately half of which falling in the 95% listener action zone (refer to Figure 5.21).

Figure 5.21: Ikwekwezi FM Fluctuation Distribution



Individual segments frequently had loudness values in the single digits, and there were a number of ads and links that measured in the -8LUFS region. The overall loudness for music segments measured -11.8LUFS. The overall loudness range values for specific programme types also varied considerably from the hyper-compressed music average of 2.2LU, to the exceptionally wide 15.1LU for advertisements. Figure 5.22 also shows that there is a median close to 6LU in the distribution of the loudness ranges of individual segments, but with a number of outliers both on the very narrow and wide loudness range sides.

Figure 5.22: Ikwewezi FM Loudness Range Distribution



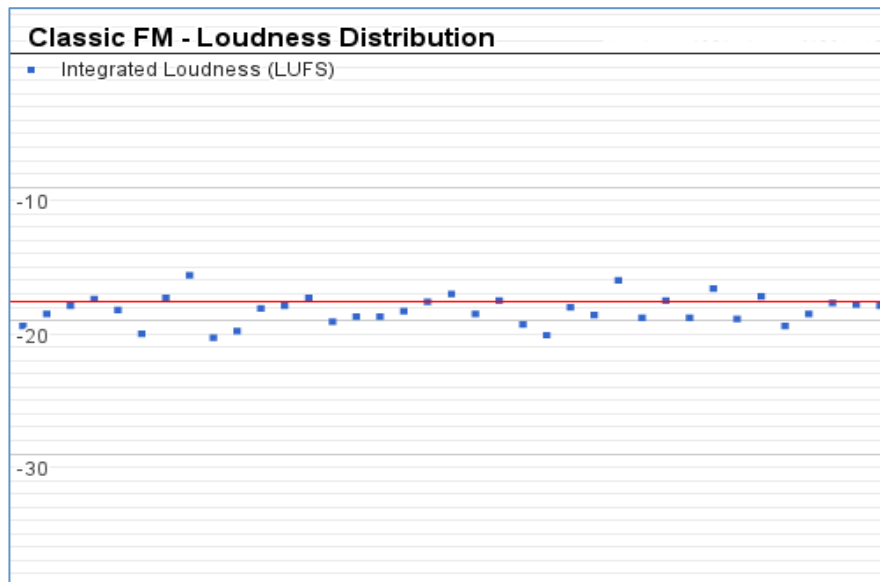
While this experiment was not set up to measure the percentage of maximum peak deviation of an FM signal, or any potential over-modulation, it would be highly surprising if there are not instances where Ikwewezi FM has an over-modulated signal outside the permitted frequency range of their station, and encroaching on adjacent stations. In terms of loudness measurement, there were a number of true peak readings exceeding 0dBFS, probably caused by inter-sample peaking – the highest of which at +2.5dBFS (dBTP). General station operation was poor with very long silences, abrupt stops, distortion on some segments and audible operation (button noise, movement in the studio, etc.).

Even if most other radio stations employed loudness normalisation, and had station loudness values within listener tolerance limits, having a station as loud as Ikwewezi as part of the offering would almost always cause problematic inter-station fluctuations. This notion is supported by the results from the ‘zap-test’ simulations in the next part of this chapter.

5.2.13 Classic FM

Classic FM is included in this section as the station delivered some results other than expected. Classical music is characterised for having wide to extremely wide loudness range values. This naturally presents some challenges in the FM broadcast environment, and some level of dynamics processing will always be required for the content to fit the medium. However, Classic FM employs dynamics processing to such an extent that the station loudness measured to be exactly the same is 5FM (-18.5LUFS), despite vastly different genres (refer to Figure 5.23). The station loudness range measured only 5.8LU, which is extremely narrow for classical music. This does enable ‘hearing more content’ while commuting, but noticeably affects the music when reproduced on a full range system in a domestic environment.

Figure 5.23: Classic FM loudness distribution of adjacent programme segments.



5.3 Zap Tests

Two different sets of zap tests were performed during data capture. The first is a test involving the top 10 radio stations, as from the previous section. The test involves live data capture where it is simulated as if a listener is ‘zapping’ through each station and waiting approximately 10 seconds on a station before moving to the next. This tests a real life scenario where inter-station loudness fluctuations may occur. Thirteen separate zap tests were conducted at different times of the day for the top 10 radio stations. The second test set involved the same simulation, but now including 10 sets of all 29 radio stations with strong reception at the testing high-site (full list with frequencies given in Table 4 in section 3.2.3).

The purpose of this extended test was twofold: firstly, it provides a legitimate scenario simulating someone tuning sequentially through all the radio frequencies, from station to station, looking for a station to listen to. Secondly, it will be tested whether zap testing using short radio clips could be aggregated to give a robust integrated station loudness by comparing their integrated loudness values to those of the top 10 stations for which much more comprehensive data was captured. It is noticeable on Table 4 that the following stations were included twice in this test simulation: Thobela FM, Motsweding, Metro FM, Talk Radio 702, Ikwekwezi FM, Ukhozi FM and SAfm. This is because these stations broadcast at different frequencies, and both frequencies were received clearly at the test site.

Figure 5.24: Integrated loudness from zap test – all stations

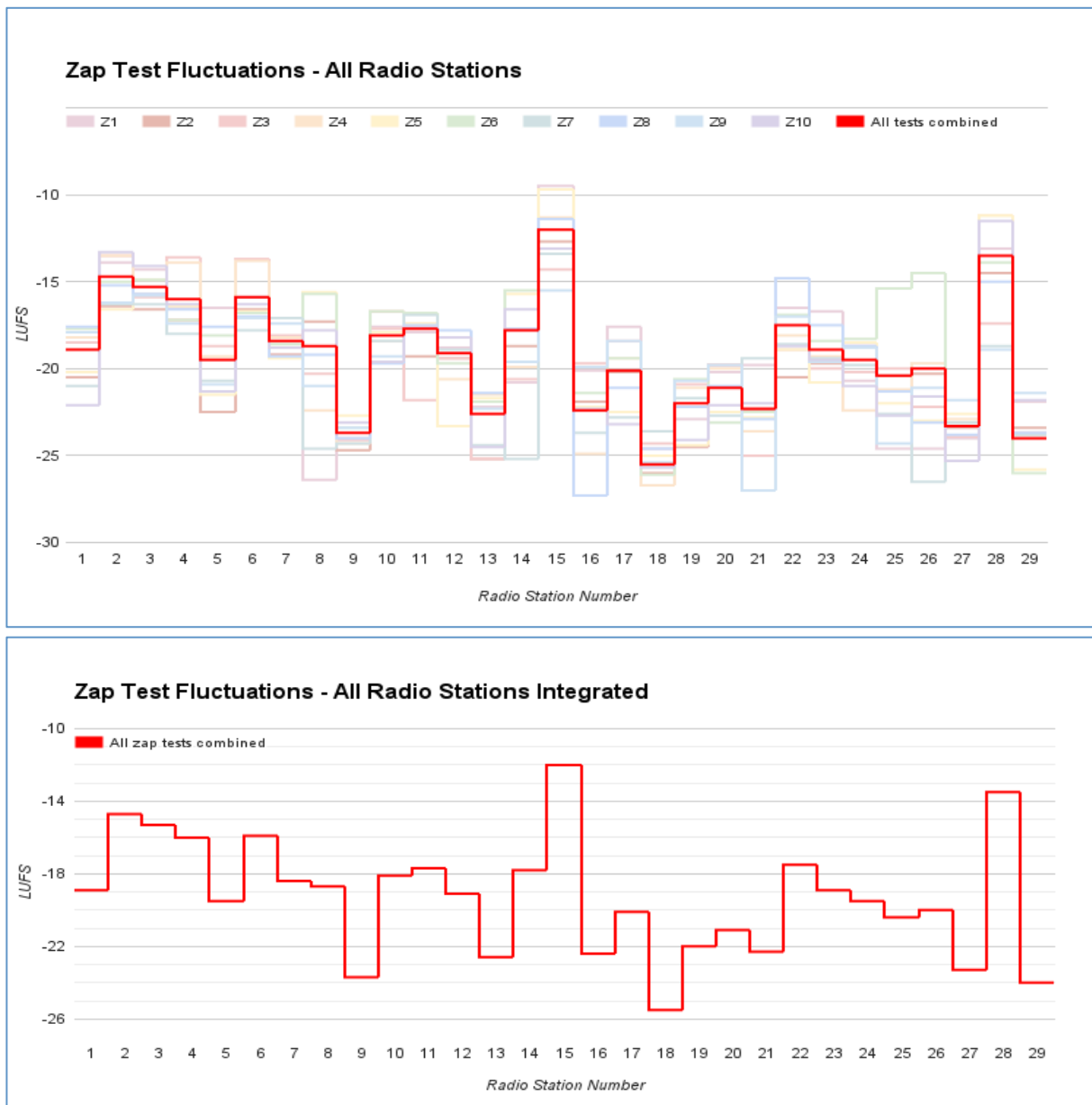


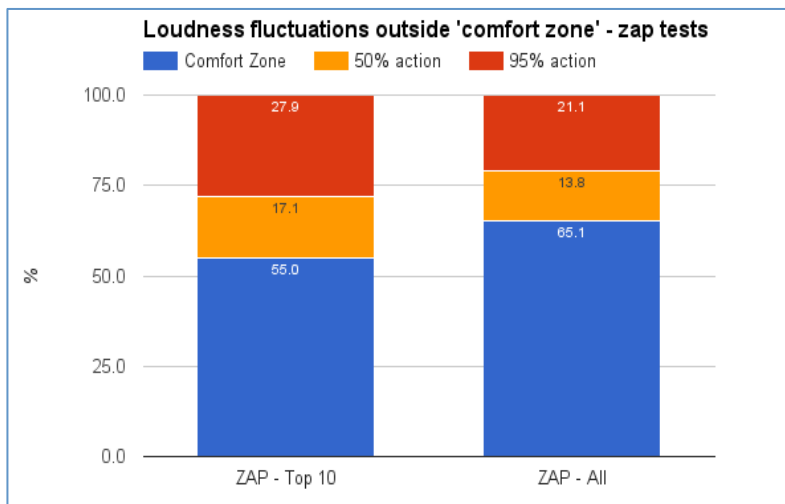
Figure 5.24 is a graphical representation of the test set involving all radio stations. The bottom half of the picture shows the integrated station loudness of each adjacent station based on 10 individual zap tests. The top half of the graph shows each individual zap test in the lighter colours, and the combined integrated result in bright red. This graph is very good indication of the presence of inter-station loudness fluctuations outside the listening ‘comfort zone’. Clearly visible in this graph are the two instances of Ikwekwezi FM (station 15 and 28), with the latter causing especially large fluctuations as it is adjacent by transmission frequency on both sides to stations with below average loudness (Talk Radio 702 and Lotus FM respectively).

Table 8: Average fluctuation magnitude for zap tests – all radio stations

	Station no.	Zap loudness	Average jump		Station no.	Zap loudness	Average jump
Thobela	1	-18.9	+5.1*	R2000	16	-22.4	-10.4
Lesedi	2	-14.7	+4.2	5FM	17	-20.1	+2.3
Ligwalagwala	3	-15.3	-0.6	Power	18	-25.5	-5.4
Motsweding	4	-16	-0.7	YFM	19	-22	+3.5
Thobela	5	-19.5	-3.5	Phalaphala	20	-21.1	+0.9
Motsweding	6	-15.9	+3.6	RSG	21	-22.3	-1.2
Ukhozi	7	-18.4	-2.5	Ukhozi	22	-17.5	+4.8
Metro	8	-18.7	-0.3	Classic	23	-18.9	-1.4
702	9	-23.7	-5	Pretoria	24	-19.5	-0.6
UW	10	-18.1	+5.6	SAfm	25	-20.4	-0.9
Jacaranda	11	-17.7	+0.4	SAfm	26	-20	+0.4
Highveld	12	-19.1	-1.4	702	27	-23.3	-3.3
MM	13	-22.6	-3.5	Ikwekwezi	28	-13.5	+9.8
Metro	14	-17.8	+4.8	Lotus	29	-24	-10.5
Ikwekwezi	15	-12	+5.8	*This jump is based on a cycle back from Lotus to Thobela. Assumes FM only seeking.			

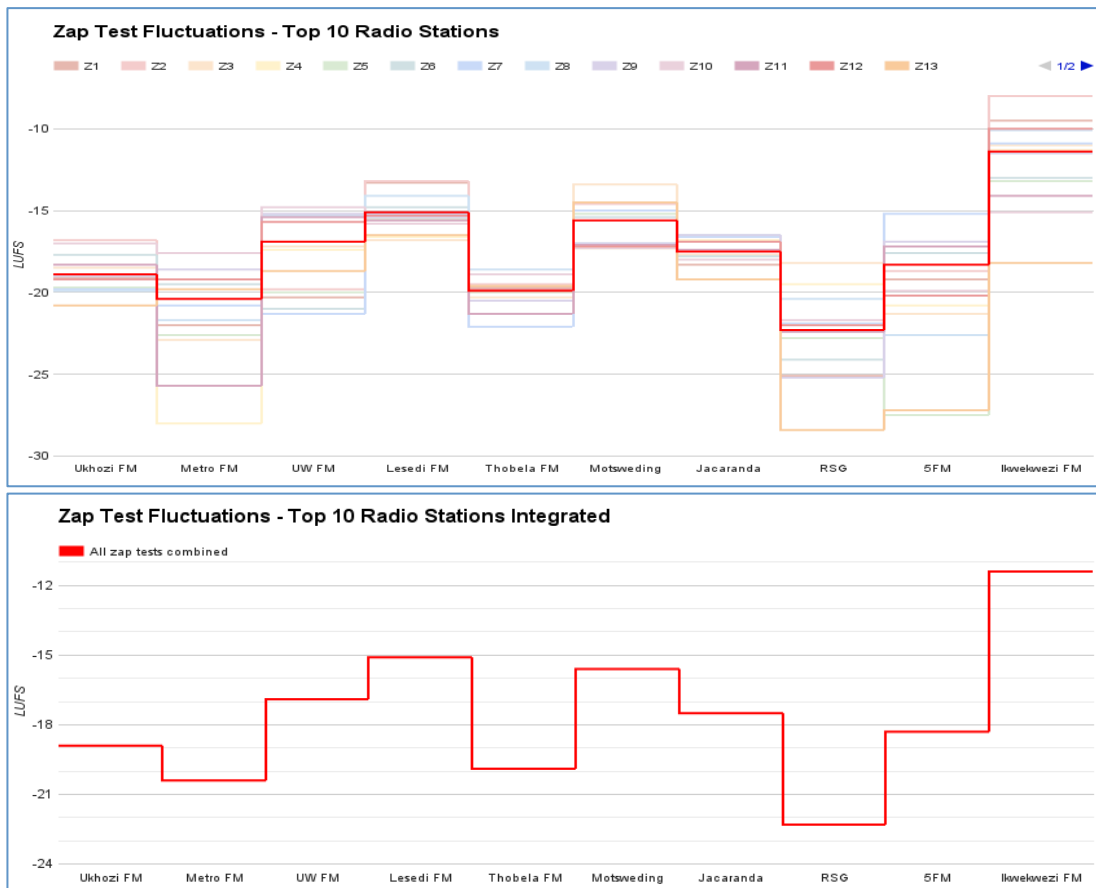
Each individual ‘zap’ was then also treated as adjacent segments, similar to those in section 5.2, but now occurring as adjacent block on different stations as presented in Figure 5.24 and Table 8. Analysing this fluctuation data showed that nearly 35% of all these fluctuations were outside the listening ‘comfort zone’, of which more than 20% were in the 95% listener action zone, as presented in Figure 5.25 below. The same methodology was employed, but this time zap tests only included the top 10 most popular radio stations, as with the rest of this study, and 13 individual zap tests were conducted across various time regions in the day.

Figure 5.25: Proportion of fluctuations outside listener ‘comfort zone’ – zap tests



An even greater proportion of problematic adjacent ‘zaps’ occurred in this test set, with 45% of adjacent blocks falling outside the listening ‘comfort zone’ as given in Figure 5.25 with individual results in Figure 5.26 below. Similar to Figure 5.24, Ikwekwezi FM causes large fluctuations. Notice also how the distributions of individual zap tests per station in the top half corresponds to the size of station loudness ranges from the previous section.

Figure 5.26: Integrated loudness from zap test – top 10 stations

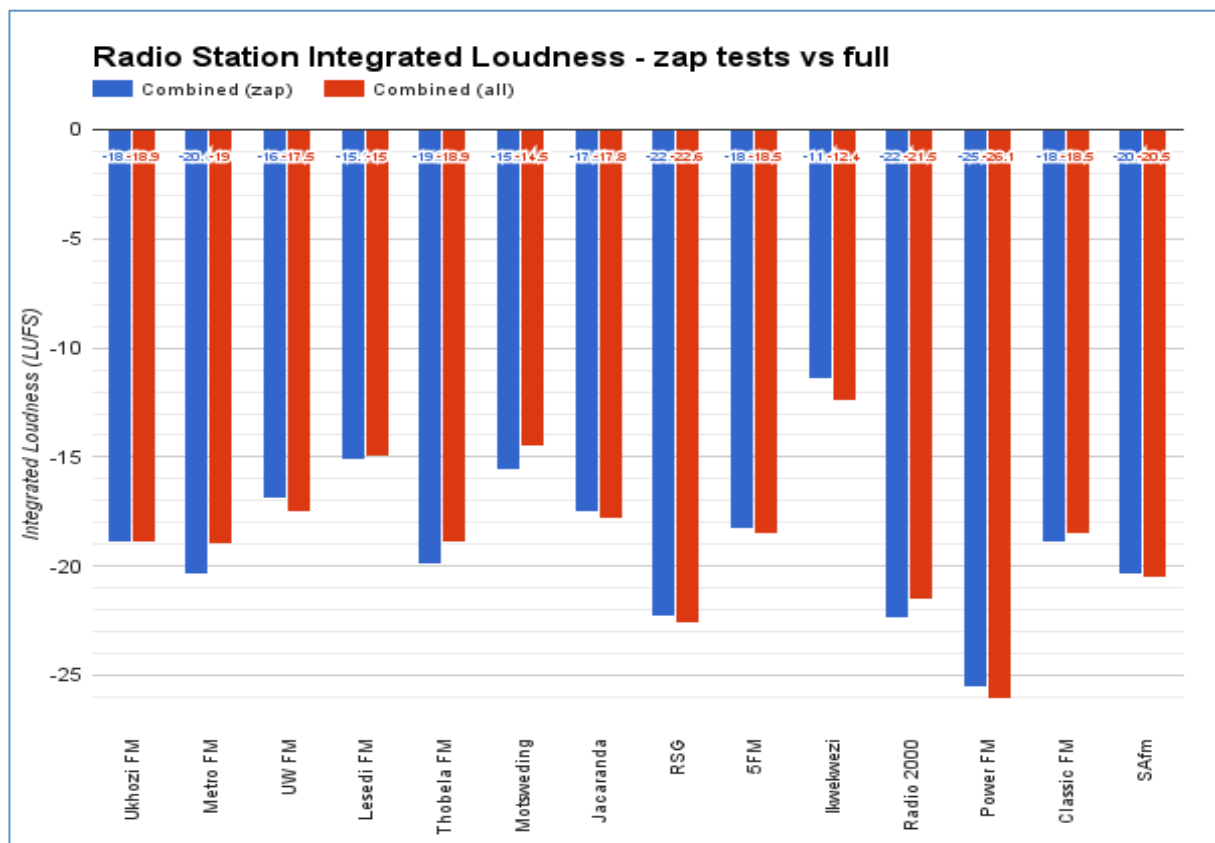


The results from all zap tests clearly show the overwhelming presence of inter-station loudness fluctuations outside the listening ‘comfort zone’. The next section will analyse whether zap test methodology could be employed to extrapolate the station loudness of all surveyed stations by comparing results to the much larger dataset from the previous section.

5.4 Zap Testing as efficient methodology for determining station loudness

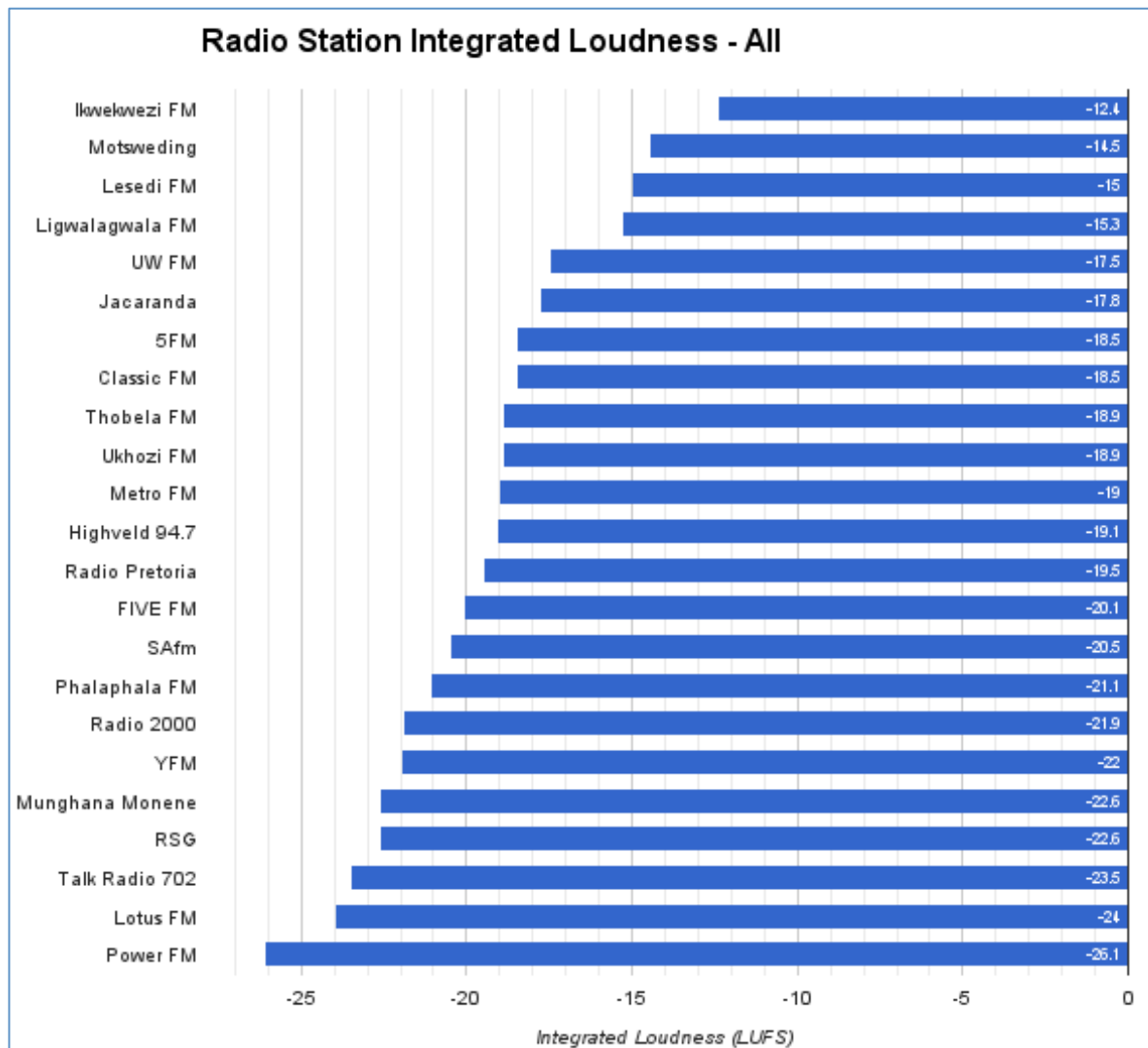
In the full testing methodology employed by this study, more than 100 individual programmes were logged per station for comprehensive loudness data. This resulted in the equivalent of 2 - 3 hours of data per station and allowed for analysis between different programme types in each station. However, if only the integration station loudness needs to be determined, it would be much more efficient if this result could be based on 10-second zap tests rather than multiple full programme segments per station. This allows for testing a large amount of stations while still performing the tests at the same time of day. Refer to Figure 5.27 below, for a comparison of integrated station loudness values between full testing (red) and zap testing (blue). Note how every result (except for Metro FM with a 1.4LU difference) falls within the ± 1 LU tolerance amounts described in EBU R128.

Figure 5.27: Integrated station loudness comparison between zap testing and full testing



The Metro FM result falls outside the basic $\pm 1\text{LU}$ tolerance, as the station itself has a high loudness range value of nearly 11LU , and despite this result, probabilities of inter-station fluctuations occurring based on data from the zap testing methodology will mostly be unaffected as the fluctuation ‘comfort zone’ spans an asymmetrical 8LU . Based on this comparison, the author can now present the overall station loudness of all 23 radio stations measured, ranked from loudest to softest, in Figure 5.28. These values will have accuracies within the tolerances of the EBU R128 recommendation. Notice that the difference between the softest and loudest stations is a staggering 13.7LU , and the median loudness between all stations is approximately -19LUFS .

Figure 5.28: Overall loudness ranking of all 23 radio stations available at testing high-site



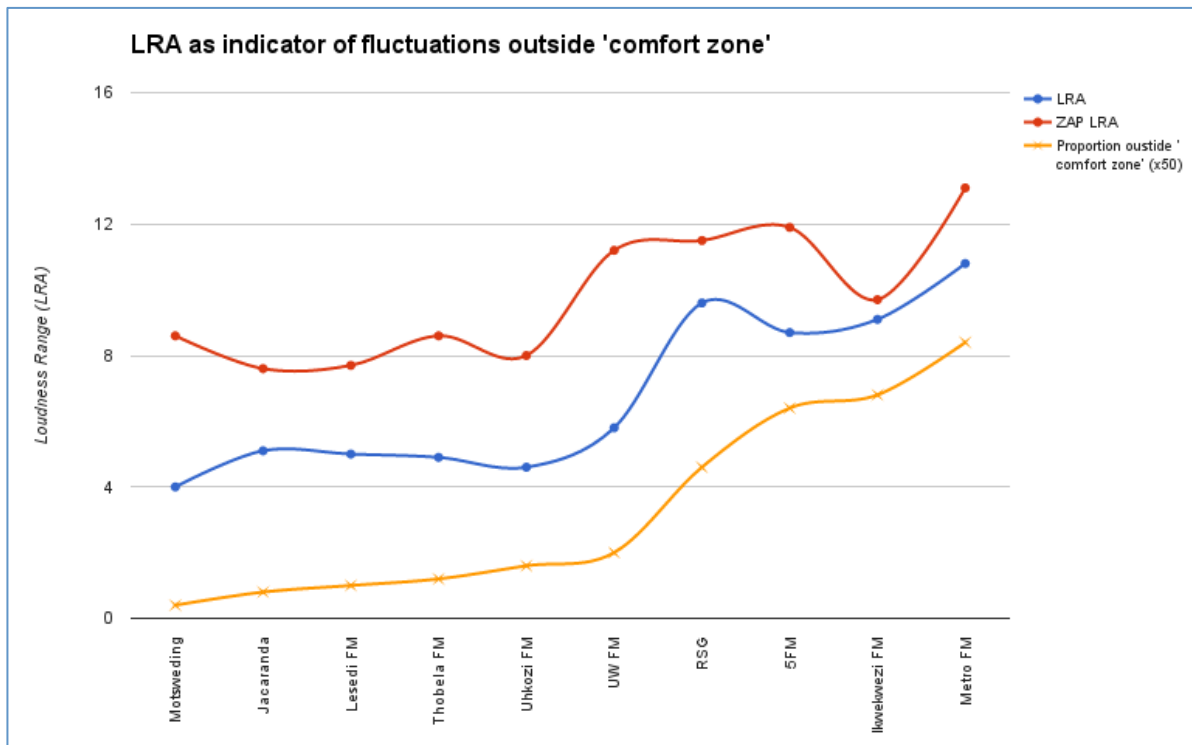
A ranking such as this, of South African radio station overall integrated station loudness has never before been released publicly. This section showed that zap test methodology can be used to extrapolate station loudness, the next section will similarly analyse the possible

correlation between overall station loudness range values and the proportion of problematic loudness fluctuations measured within each station.

5.5 Loudness Range values as fluctuation indicators

Loudness range is meant as a descriptive measure of the difference between the average loudest and average softest parts of a single audio programme longer than at least 30 seconds (see sections 2.2.3 and 2.3.2 for further details). The measure is meant to be used in conjunction with integrated loudness values, and can be useful to determine platform suitability, and to indicate content with a wide loudness range. This measure has generally not been used to describe an entire station, but the author noticed a possible correlation between the proportion of problematic loudness fluctuations and an increase in station loudness range values. This correlation is presented in Figure 5.29 below. The graph indicates both the full LRA values based on comprehensive data logging, and the LRA values based on the zap testing methodology. The values indicated by the yellow line multiplied by 50 gives a percentage of all fluctuations falling outside the listening ‘comfort zone’. Neither shows a very strong correlation, but there is a general trend in that higher amounts of fluctuations correspond to a higher LRA value – especially on the comprehensive dataset.

Figure 5.29: Loudness range as indicator of problematic fluctuations



It can be reasoned that in radio stations that broadcast content with a higher loudness range, but normalise content based on integrated loudness, would not exhibit this correlation. Similarly, a station that has expert (manual) control over the content transmitted may also exhibit a wide loudness range, if programmes overall get louder and softer, but adjacent programming is kept within tolerance limits. Further research would be required to investigate this correlation.

5.6 Conclusion

This chapter explored the two subsidiary research questions in detail, with specific focus on intra-station loudness fluctuations in radio broadcasts. An overview was given for each individual radio station, including additional observations made during data capture. All of the top 10 radio stations exhibited some loudness fluctuation outside the listening ‘comfort zone’, with Metro FM the most and Motswedding the least. The results from simulated inter-station fluctuation testing (or ‘zap testing’) was presented, showing that 35% of inter-station changes between all available radio stations resulted in problematic fluctuation, and 45% of inter-station changes between the top 10 stations resulted in problematic fluctuations. Zap testing was also presented as an efficient methodology for determining station loudness, and showed a very strong correlation when comparing the results to station loudness testing based on the full dataset. This enabled the extrapolation of station loudness data of all 23 radio stations available at the test site. Lastly, station loudness range values were suggested as a possible indicator for the proportion of problematic loudness fluctuations in non-loudness normalised stations. Some correlation was found. The next and final chapter of this dissertation will present a summary of conclusions and recommendations of this study.

Chapter 6

Conclusion and Recommendations

This chapter will present a summary of all the findings from the previous two chapters and present it in the structure of the study's research questions. It will explore how production houses and broadcasters can employ loudness normalisation to prevent problematic loudness fluctuations, including the benefits and challenges of making such an operational paradigm shift. Lastly, a brief section highlighting proposed future research in the subject field is provided.

6.1 Summary of findings

The first part of this summary answers the main research question of the study, “*does South African free-to-air broadcast audio contain loudness fluctuations of a magnitude that exceeds previously determined limits of listening comfort?*” The answers to this will be summarised separately for television and radio audio.

South African free to air television stations broadcast over radio frequencies⁵ exhibit inter-station loudness fluctuations outside the viewer ‘comfort zone’. When switching between these stations there is a high probability of experiencing inter-station loudness fluctuations outside the +3LU -5LU asymmetrical listening ‘comfort zone’, especially switching between SABC 2 and SABC 3. However, due to aggressive dynamics processing prior to broadcast, none of these television stations exhibit any intra-station fluctuation outside the ‘comfort zone’ – neither by adjacent programming block, genre, day of week nor time of day. The dynamics processing does have noticeable artefacts, but these effects are beyond what the metrics of this study are able to measure. A reader that finds it hard to believe that these stations do not contain intra-station fluctuations should keep in mind that this result cannot be compared to digital broadcast streams (even for the same stations).

Due to the large number of free to air radio stations, this study focused on the top 10 most popular stations based on the Radio Audience Measurement Survey, that collectively covers

⁵ As opposed to the same free-to-air broadcasts made available through digital video broadcast (DVB) providers such as StarSat and DSTV.

the majority of the South African radio audience. South African free to air radio stations contained a large amount of loudness fluctuations exceeding previously determined limits of listening comfort. They presented vast differences between overall station loudness and to a lesser extent between programme types within stations. However, they did not present any loudness trends when comparing different times of the day, or comparing weekday to weekend programming. Across the board, advertisements were not the loudest, but rather links (interstitials) or music segments. On all stations, talking was broadcast the softest, followed by advertisements.

The first subsidiary research question was, *“what is the extent of loudness fluctuations in South African broadcast audio, both within and between channels, in television and radio respectively, and do they exceed the previously determined limits of listening ‘comfort zone’?”* Again, this will be answered separately for television and radio audio.

For television broadcast, the extent of inter-station fluctuations is a probability based on the combination of the station’s overall integrated loudness and loudness range. Measurements showed the SABC 1 to be at -18.9LUFS station loudness and 4.6LU station loudness range, SABC 2 at -14.8LUFS station loudness and 4.9LU station loudness range, SABC 3 at -20.2LUFS station loudness and 5.5LU station loudness range, and e.tv at -16.3LUFS station loudness and 6.8LU station loudness range. This shows that the probability will always exist for intra-station loudness fluctuations outside the listener ‘comfort zone’.

For radio broadcasts, larger probabilities exist for inter-station fluctuations, as the difference between the softest and loudest stations measured is high, at 13.7LU, and the median loudness between all stations is approximately -19LUFS. Ikwewezi FM is by far the loudest station measured at -12.4LUFS integrated station loudness, with the softest being Lotus FM at -26.1LUFS. Inter-station fluctuations were further tested using ‘zap test’ methodology to simulate actual inter-station fluctuation use-cases. In these tests, 45% of adjacent ‘zap’ segments contained problematic fluctuations between the top 10 radio stations, and nearly 35% of adjacent ‘zap’ segments contained problematic fluctuations when testing all 29 stations available at the testing site. In terms of intra-station fluctuations, all of the top 10 radio stations exhibited loudness fluctuations outside the ‘comfort zone’, with Motsweding the least at only 2% of total adjacent programmes, and Metro FM the most with 42% of adjacent programmes.

It was also found that ‘zap test’ methodology (measuring approximately 10 second segments while cycling through all stations) provided for a robust measure of station loudness with as little as 10 zap tests performed. When comparing the station loudness values from full testing to those using ‘zap test’ methodology, there was a strong correlation, and results fell within the $\pm 1\text{LU}$ tolerance as specified by the EBU R128 recommendation used in the study. This enabled the extrapolation of station loudness, and accordingly the ranking of all 23 free to air radio stations available at the testing high-site.

The second subsidiary research question was, *“does the broadcast audio programming contain a Loudness Range that is appropriate for the intended platform, listening environment and any fluctuations that exceed previously determined limits of Dynamic Range Tolerance of the listener?”*

With some exceptions, the answer to this question is generally, yes, the vast majority of broadcast audio measured would be appropriate for the intended platform. In fact, most television and radio stations have a loudness range far narrower than the threshold for household viewing or listening. Some of the radio stations (most notably Metro FM, 5FM, RSG and Ikwewezi FM) frequently contained loudness ranges on the segment level that was too wide to be appropriate for listening in transit, or on portable devices. Generally, almost all content was appropriate for a domestic setting. However, it was also explored whether station loudness range values could be an indicator for the probability of problematic fluctuations. Small loudness range values were observed to correspond with highly compressed and peak-limited content, especially values related to individual programme segments and also on station level. The study found some level of positive correlation between the station loudness range values and an increase in the total proportion of problematic loudness fluctuations in the station.

6.2 Recommendations for broadcasters and content producers

It is recommended that all South African broadcast audio producers and distributors begin the process to make the shift from peak normalisation to loudness normalisation, as this would prevent problematic loudness fluctuations. The most sensible standard to follow would be the EBU R128 recommendation, for the following three reasons:

1. It is based on open, international standards
2. The SABC is a member of the EBU and already uses previous EBU recommendations (such as EBU R68)

3. The PLOUD group responsible for the R128 recommendation consists of a highly active, multi-national group that is constantly improving the finer details of the recommendation, including practical information on implementation.

For content producers, the paradigm shift to loudness normalisation does not need to have large (or indeed any) financial implications. Loudness meters are included as standard in many of the new digital audio workstations (e.g. Steinberg Nuendo and Presonus Studio One) and non-linear video editors (e.g. Adobe Premiere Pro CC and Apple Final Cut X). It would require an adjustment in fixed monitoring levels, but most importantly, a change in mind-set, to mix using your ears rather than your eyes.

For broadcasters, the shift will require capital investment for new equipment, but most of the commercially available options have been built to integrate with existing broadcast workflows. There is little to add in terms of recommendations for production and distribution according to a loudness normalisation paradigm that has not already been covered in great detail in the EBU Tech documents 3343 and 3344 (EBU 2011d; EBU 2011e). In terms of South Africa's national broadcaster, the SABC, it would be opportune to make this transition than to do it simultaneously to the move from analogue to digital terrestrial broadcasting. Seeing as this move will already require large equipment upgrades and resource overheads, the additional expense would be small compared to the rest of the change, and much smaller than making the paradigm shift at a later stage.

The final recommendation from this paper also introduces a new term in the field, namely 'loudness advocacy'. While such advocacy is already underway, it has not been given this name. A good example of loudness advocacy is the work done by Florian Camerer, the chairman of the PLOUD group at the EBU. He actively engages institutions and individuals around the world, presenting workshops, providing training and engaging in meaningful discussion around this topic. It is essential to have individuals to champion the cause, but tertiary and other training institutions also have an important role to play in leading the charge in the paradigm shift to loudness normalisation, especially as they have the enhanced ability to stay abreast of latest academic progress in the field.

It is important that institutions focus on open, international standards to ensure interoperability and transferability of content across regions, and to avoid being locked into proprietary solutions.

6.3 Suggestions for further research

1. Additional research is required to investigate using station loudness range measures as an indicator for the proportion of problematic fluctuations within that station. Further research is also required on ‘zap test’ methodologies that include a randomised rather than sequential station order, including possible additional uses for ‘zap test’ data.
2. Future research could include additional and alternative delivery media, including Digital Audio Broadcast (DAB), Digital Video Broadcast (DVB) and delivery stream via the Internet Protocol (IP), especially when comparing loudness characteristics of exactly the same content over various transmission channels, for example the same radio station over RF, DAB and IP. There is also a need for research for the translation of the same audio content across different devices and platforms, for example the loudness characteristics and effects of loudness normalisation on the same content being viewed in a home theatre system, thing flat screen TV and a portable device. There has also been very little research on loudness normalisation for game audio, and other interactive audio-visual media.
3. iTunes music, iTunes radio, Spotify, YouTube and other online content delivery services have all introduced various forms of loudness normalisation. Further research could compare the currently loudness normalisation approaches of these services, and also extend to various forms of loudness normalisation of recorded music as a way to counteract the so-called “loudness war”.
4. Research is also required on the monitoring calibration standards for content producers, taking into account the new loudness measurement standards. This needs to include varying room sizes in the production environment, as well as problems associated with calibrating the Low Frequency Effects (LFE), or sub-woofer channel.
5. Loudness normalisation standards such as EBU R 128 is increasingly present in the South African commercial audio space, but potential research could include subjective studies about audio technicians’ experience making the paradigm shift, delivery to multiple specifications, and working with loudness meter ballistics.

List of References

- Benjamin, E. 2002. Comparison of Objective Measures of Loudness Using Audio Program Material. In *Proceedings of the 113th Audio Engineering Society Convention*, Los Angeles, California.
- Camerer, F., Grimm, E., Gross, K., Katz, B., Ludwig, B. & Lund, T. 2012. Loudness Normalization : The Future of File-Based Playback. Internet: http://music-loudness.com/PDFs/Loudness_Alliance_White_Paper_final_v1.pdf. Accessed: 24 February 2014.
- Camerer, F. 2010. On The Way to Loudness Nirvana - Audio Levelling with EBU R 128. *EBU Technical Review*, Q3: 1–9.
- Couling, J. 1997. *TV Loudness: Time for a New Approach?* London: Dolby Laboratories.
- Devine, K. 2013. Imperfect Sound Forever: Loudness Wars, Listening Formations and the History of Sound Reproduction. *Popular Music*, 32 (02): 159–176.
- DOC. 2014. Digital Migration. *Department of Communications (DOC)*. Internet: <http://www.doc.gov.za/key-programmes/digital-terrestrial-television-dtt.html>. Accessed: 7 July 2014.
- DSTV. 2014. Material Requirements & Deadlines. *DSTV Media Sales*. Internet: <http://dstvmediasales.com/wp-content/uploads/2014/02/DMS-Material-Requirements-Deadlines-Jan-2014.pdf>. Accessed: 16 March 2014.
- EBU. 2000. Recommendation R 68 - Alignment Level in Digital Audio Production Equipment and in Digital Audio Recorders. Geneva: European Broadcasting Union (EBU).
- EBU. 2009. *10 Things You Need to Know About Loudness*. Geneva: European Broadcasting Union (EBU).
- EBU. 2011a. Recommendation R 128 - Loudness Normalisation and Permitted Maximum Level of Audio Signals. Geneva: European Broadcasting Union (EBU).
- EBU. 2011b. Tech 3341 - Loudness Metering: “EBU Mode” Metering to Supplement Loudness Normalisation in Accordance with EBU R 128. Geneva: European Broadcasting Union (EBU).
- EBU. 2011c. Tech 3342 - Loudness Range: A Measure to Supplement Loudness Normalisation in Accordance with EBU R 128. Geneva: European Broadcasting Union (EBU).
- EBU. 2011d. Tech 3343 - Practical Guidelines for Production and Implementation in Accordance with EBU R 128. Geneva: European Broadcasting Union (EBU).

EBU. 2011e. Tech 3344 - Practical Guidelines for Distribution Systems in Accordance with EBU R 128. Geneva: European Broadcasting Union (EBU).

Emmett, J. 2003. Audio Levels in the New World of Digital Systems. *EBU Technical Review*, January: 1–5.

FACTS. 2002. *Investigation Into Loud Advertisements on Television*. Sydney: Federation of Australian Commercial Television Stations (FACTS).

Fletcher, H. & Munson, W.A. 1933. Loudness, Its Definition, Measurement and Calculation. *Bell System Technical Journal*, 12 (4): 377–430.

Gallagher, M. 2009. *The Music Tech Dictionary: A Glossary of Audio-Related Terms and Technologies*. Boston: Cengage Learning.

Grimm, E., Skovenborg, E. & Spikofski, G. 2010. Determining an Optimal Gated Loudness Measurement for TV Sound Normalization. *Audio Engineering Society Convention*, 128: 1–12.

Huber, M. & Runstein, R. 2014. *Modern Recording Techniques* 8th ed. Burlington: Focal Press.

IEC. 1991. IEC Standard 60268. Sound System Equipment - Part 10: Peak Programme Level Meters. Geneva: International Electrotechnical Commission (IEC).

ISO. 2003. *ISO Standard 226. Normal Equal-Loudness-Level Contours*. Geneva: International Organization for Standardization (ISO).

ITU. 1988. Lining-up and Monitoring an International Sound-Programme Connection. Geneva: International Telecommunications Union (ITU).

ITU. 2012. ITU-R BS.1770-3 Algorithms to Measure Audio Programme Loudness and True-Peak Audio Level. 3. Geneva: International Telecommunications Union (ITU).

Katz, B. 2007. *Mastering Audio: The Art and the Science* 2nd ed. Burlington: Focal Press.

Kruger, K. 2013. Technical Delivery Standards for SD and HD (e.tv). Internet: [http://www.etv.co.za/sites/etv.co.za/files/files/Technical Delivery Standards 2013 - 2014.pdf](http://www.etv.co.za/sites/etv.co.za/files/files/Technical%20Delivery%20Standards%202013%20-%202014.pdf). Accessed: 7 July 2014.

Kuipers, F., Kooij, R., de Vleeschauwer, D. & Brunnström, K. 2010. Techniques for Measuring Quality of Experience. In *Wired/wireless Internet Communications*, Berlin: Springer.

- Lindström, M. 2013. Evaluation of Momentary Loudness Metering Ballistics for Achieving Loudness Targets and Stable Loudness Levels in Live Broadcast. Bachelor of Arts Audio Engineering. Luleå University of Technology.
- Lund, T. 2006a. Control of Loudness in Digital TV. *Proceedings of the NAB Convention, April 2006*: 57–65.
- Lund, T. 2006b. Stop Counting Samples. *Audio Engineering Society Convention*, 121: 1 – 11.
- Lund, T. 2008. Inter-Program Level Jumps in Broadcast. *Broadcast Asia*: 1–7.
- Lund, T. 2011. ITU-R BS. 1770 Revisited. *Proceedings of the NAB Convention, April 2011*: 1–11.
- Lund, T. 2013. Audio for Mobile TV, iPad and iPod. In *National Association of Broadcasters (NAB) Convention*, Las Vegas.
- Mahomed, Z. & Shabangu, F. 2011. Media Solutions Technical Standards. Internet: <http://images.supersport.co.za/SMSTechnicalStandardsOctober2011v5.pdf>. Accessed: 6 June 2014.
- Miyasaka, E. & Kamada, T. 2006. Effects of TV Commercials Inserted in Main Programs on Viewers. *Bulletin of Faculty of Environmental and Information Studies, Musashi Institute of Technology*, 8: 172–183.
- Moore, B., Glasberg, B. & Stone, M. 2003. Why Are Commercials so Loud? *Journal of the Audio Engineering Society*, 51 (12): 11–15.
- Mouton, J. 2001. How to Success in Your Master's & Doctoral Studies. Pretoria: Van Schaik.
- Norcross, S., Lavoie, M. & Thibault, L. 2011. *A Status Report on Loudness Control Technologies and Standardization for Broadcasting*. Ottawa: Communications Research Centre Canada (CRC).
- OED. 1993. Dynamic Range. In *Oxford English Dictionary (OED)*, Oxford: Oxford University Press.
- Orban. 2014. Orban Loudness Meter. Internet: <http://www.orban.com/meter/>. Accessed: 6 June 2014.
- PMF. 2015. Our Aim. *Pleasurize Music Foundation (PMF)*. Internet: <http://www.pleasurizemusic.com>. Accessed: 23 May 2015.
- Riedmiller, J., Lyman, S. & Robinson, C. 2003. Intelligent Program Loudness Measurement and Control: What Satisfies Listeners? *Audio Engineering Society Convention*, 115: 1–16.

Robinson, D. & Dadson, R. 1956. A Re-Determination of the Equal-Loudness Relations for Pure Tones. *British Journal of Applied Physics*, 7 (5): 166–181.

SAARF. 2014. Website of the South African Audience Research Foundation (SAARF). Internet: <http://www.saarf.co.za/>. Accessed: 7 July 2014.

SAARF. 2015a. Radio Audience Measurement Survey (RAMS) June 2015. *South African Audience Research Foundation (SAARF)*. Internet: http://www.saarf.co.za/rams-commercial/2015/RAMS_JUNE_2015-Audience_National.pdf. Accessed: 3 June 2015.

SAARF. 2015b. RAMS Methodology. Internet: <http://www.saarf.co.za/RAMS/rams-methology.asp>. Accessed: 20 June 2015.

SABC. 2003. SABC Policy and Procedures for the Procurement of Local Television Programmes. Johannesburg: South African Broadcasting Corporation (SABC).

SABC. 2004. *SABC Technical Standards for Television Programme Delivery*. Johannesburg: South African Broadcasting Corporation (SABC).

Sack, M.C. 2011. *Loudness and Auditory Masking for Mobile TV*. Masters of Media Informatics. University of Vienna.

Sengpiel Audio. 2005. Normal Equal-Loudness-Level Contours - ISO 226 : 2003 Acoustics International Organization for Standardization (ISO) 2nd Edition. Internet: <http://www.sengpielaudio.com/Acoustics226-2003.pdf>. Accessed: 13 June 2015.

Skovborg, E. & Lund, T. 2008. Loudness Descriptors to Characterize Programs and Music Tracks. *Audio Engineering Society Convention*, 125: 1–9.

Skovborg, E. & Lund, T. 2009. Loudness Descriptors to Characterize Wide Loudness-Range Material. *Audio Engineering Society Convention*, 127: 1–12.

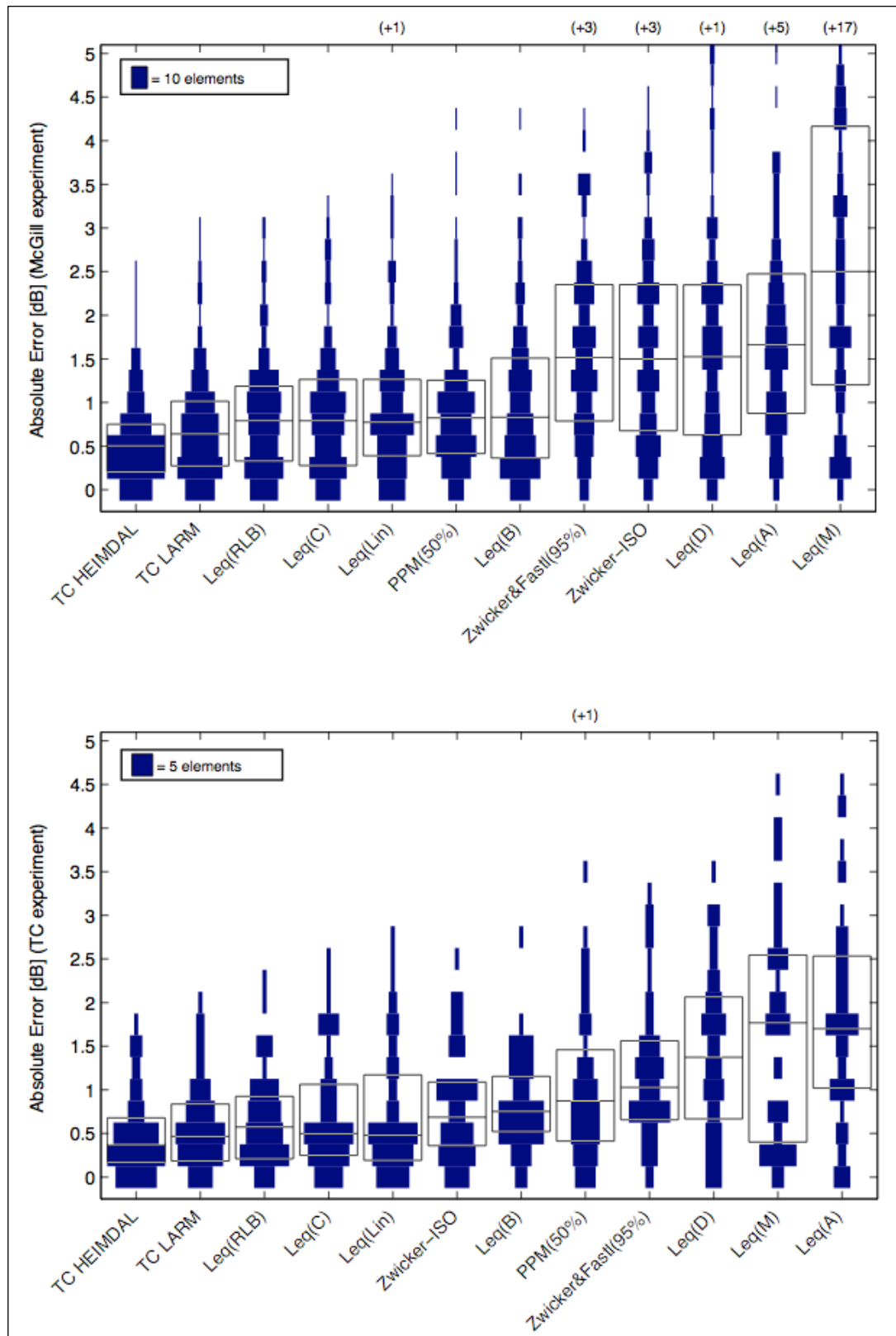
Skovborg, E. & Nielsen, S. 2004a. Evaluation of Designs for Loudness-Matching Experiments. In *Proceedings of the International Conference “Subjective and Objective Assessment of Sound” (SOAS)*, Poznan, Poland.

Skovborg, E. & Nielsen, S. 2004b. Evaluation of Different Loudness Models with Music and Speech Material. *Audio Engineering Society Convention*, 117: 1 – 34.

- Skovborg, E. & Nielsen, S. 2007. Real-Time Visualisations of Loudness Along Different Time Scales. In *10th International Conference on Digital Audio Effects (DAFx)*, Bordeaux.
- Soulodre, G. 2004. Evaluation of Objective Loudness Meters. In *Proceedings of the 116th Audio Engineering Society Convention*, New York.
- Spikofski, G. & Klar, S. 2004. Levelling and Loudness-in Radio and Television Broadcasting. *EBU Technical Review*, January: 1–12.
- Steinberg GmbH. 2015. The Whole World of Audio Post. Nuendo 6.5. Internet: https://www.steinberg.net/en/products/nuendo_range/nuendo/whats_new/whats_new_in_nuendo_65.html. Accessed: 12 June 2015.
- Taylor, R.W. & Martens, W.L. 2014. Hyper-Compression in Music Production: Listener Preferences on Dynamic Range. *Audio Engineering Society Convention*, 136: 1–10.
- Vickers, E. 2010a. Metrics for Quantifying Loudness and Dynamics. *Audio Engineering Society Convention*, 129: 1–7.
- Vickers, E. 2010b. The Loudness War: Background, Speculation and Recommendations. *Audio Engineering Society Convention*, 129: 1–27.
- Weymouth, D. 2012. *The Loudness War: A Game and Market Theory Analysis*. Munich: GRIN Verlag.

Appendix A – Loudness model comparisons

Figure 7.1: Distribution of absolute error for evaluated loudness models



Taken from Skovenborg & Nielsen (2004b: 21).

Appendix B – Full Radio Audience Measurement Survey (RAMS)

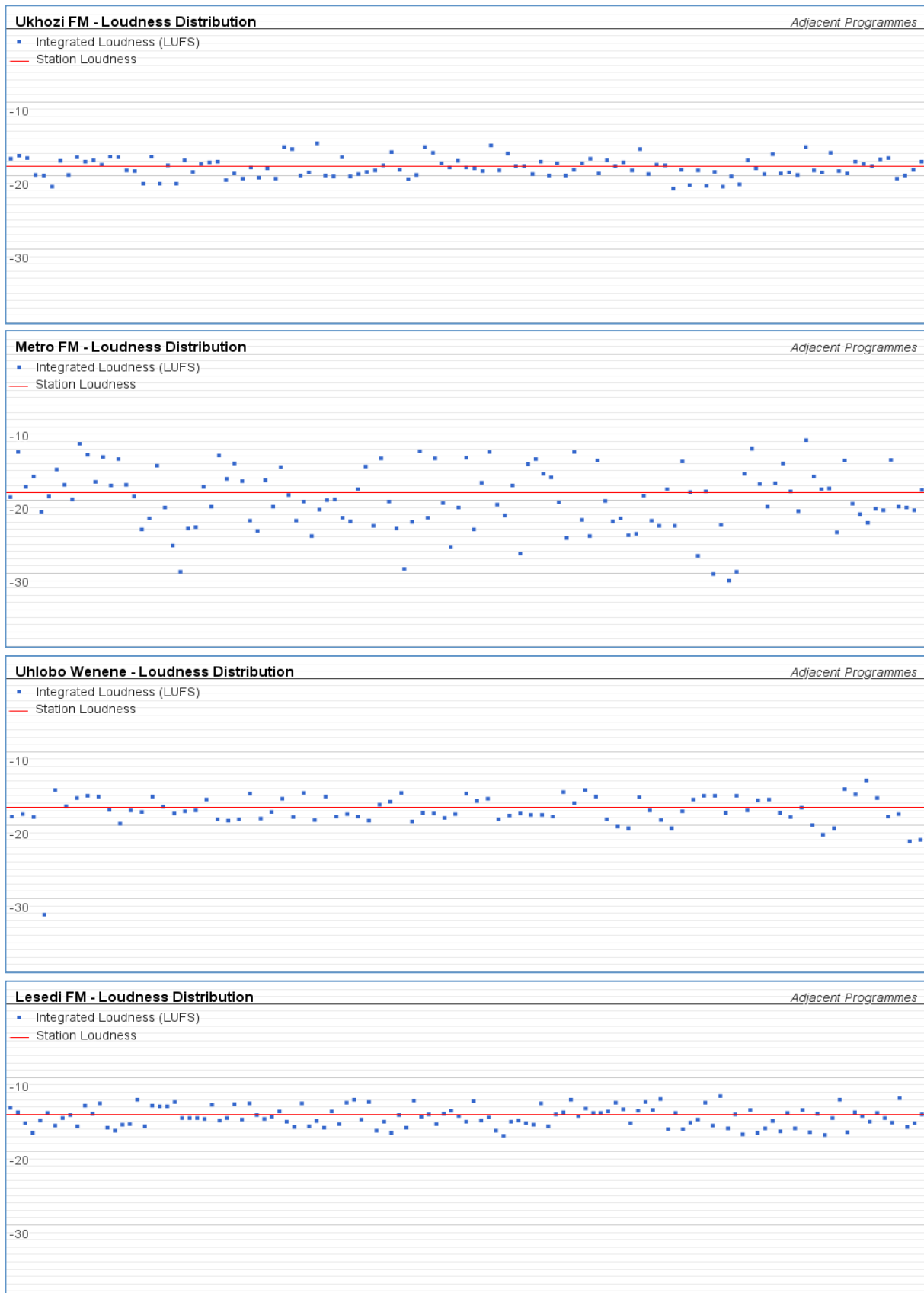
Table 9: Radio Audience Measurement Survey for May 2014 - June 2015

	MAY '14	JUNE '14	AUG '14	OCT '14	DEC '14	FEB '15	JUN '15		
STATION	%	%	%	%	%	%	%	Ave	Rank
UKHOZI FM	20.2	20.5	20.5	20.9	20.7	20.5	20	20.5	1
METROFM	17.3	17	16.9	18	17.4	17	16.9	17.2	2
UMHLOBO WENENE FM	11.9	12	12.1	11.7	11.4	11.3	12.3	11.8	3
LESEDI FM	10.9	10.7	10.5	10	9.6	9.7	9.6	10.1	4
THOBELA FM	8.8	9	8.9	9	9	9.2	7.9	8.8	5
MOTSWEDING FM	9.1	8.9	8.8	8.2	8.3	8.3	8	8.5	6
JACARANDA 94.2	5.5	5.6	5.5	5.1	5.1	5.1	5.2	5.3	7
RSG	5.1	5.1	5.1	5.1	5.2	5.1	4.9	5.1	8
GAGASI FM	5.1	5	5	5	5	5	4.4	4.9	9
SFM	5.2	5	5.3	5.1	4.6	4.5	4.6	4.9	9
IKWEKWEZI FM	4.8	5	4.7	4.8	4.9	4.7	4.5	4.8	10
EAST COAST RADIO	4.6	4.5	4.3	4.5	4.4	4.5	3.8	4.4	11
KAYAFM 95.9	4.4	4.2	4.3	4.6	4.2	3.9	4.2	4.3	12
LIGWALAGWALA FM	4.1	4.1	4.2	3.7	3.6	3.6	3.9	3.9	13
YFM	4.3	4.2	3.7	3.7	3.6	3.5	3.7	3.8	14
HIGHVELD 94.7	4	3.9	3.6	3.7	3.8	3.5	3.3	3.7	15
CAPRICORN FM	3.8	3.9	3.9	3.3	3.3	3.3	2.3	3.4	16
MUNGHANA LONENE FM	2.9	2.8	2.7	2.7	2.7	2.6	3.1	2.8	17
KFM	2.7	2.7	2.6	2.8	2.8	2.6	2.6	2.7	18
PHALAPHALA FM	2.3	2.5	2.5	2.5	2.4	2.3	2.5	2.4	19
RADIO 702	2.3	2.3	2.4	2.5	2.3	2.2	2.2	2.3	20
RADIO 2000	2.5	2.4	2.1	2.1	2.4	2.2	2.3	2.3	20
GOOD HOPE FM	1.9	1.9	1.9	1.7	1.7	1.7	1.7	1.8	21
ALGOA FM	1.7	1.7	1.6	1.7	1.8	1.7	1.6	1.7	22
HEART FM	1.7	1.8	1.7	1.6	1.5	1.6	1.7	1.7	22
SAFM	1.7	1.6	1.5	1.5	1.4	1.4	1.5	1.5	23
OFM	1.1	1.2	1.3	1.1	1.1	1.1	1.1	1.1	24
LOTUSFM	1	0.9	0.9	1	1	0.9	0.9	0.9	25
NORTH WEST FM	1.1	1	1	0.8	0.8	0.7	0.6	0.9	26
VUMA FM	0.4	0.5	0.5	0.9	0.9	0.7	0.8	0.7	27
TRUFM	0.6	0.6	0.6	0.5	0.5	0.5	0.5	0.5	28
CLASSIC FM	0.4	0.4	0.4	0.5	0.4	0.5	0.5	0.4	29
RADIO PULPIT	0.3	0.3	0.2	0.2	0.3	0.4	0.3	0.3	30
CAPE TALK	0.3	0.3	0.3	0.3	0.2	0.3	0.3	0.3	30
ALGOA FM	0.2	0.2	0.3	0.2	0.2	0.2	0.3	0.2	31
POWER 98.7	0.2	0.3	0.2	0.2	0.2	0.2	0.2	0.2	31
SMILE FM	0.1	0.1	0.1	0.1	0.1	0.2	0.2	0.1	32

Adapted from SAARF (2015a: 5).

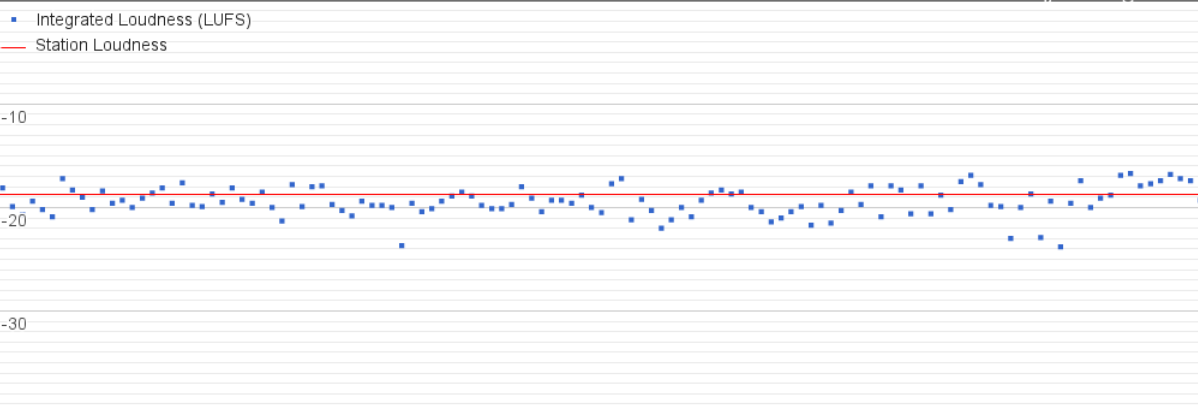
Appendix C – Full Programme Loudness Distribution Graphs

Figure 7.2: Loudness distributions of all adjacent programmes in radio stations



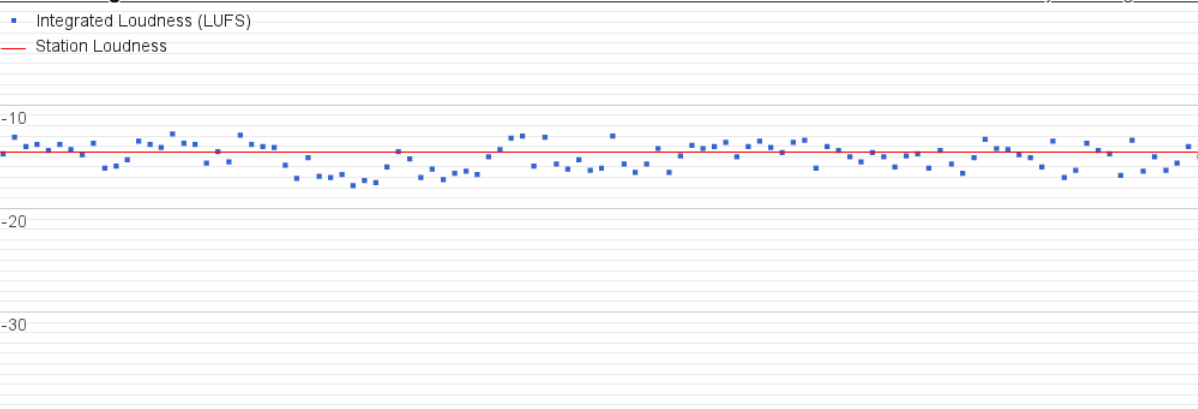
Thobela FM - Loudness Distribution

Adjacent Programmes



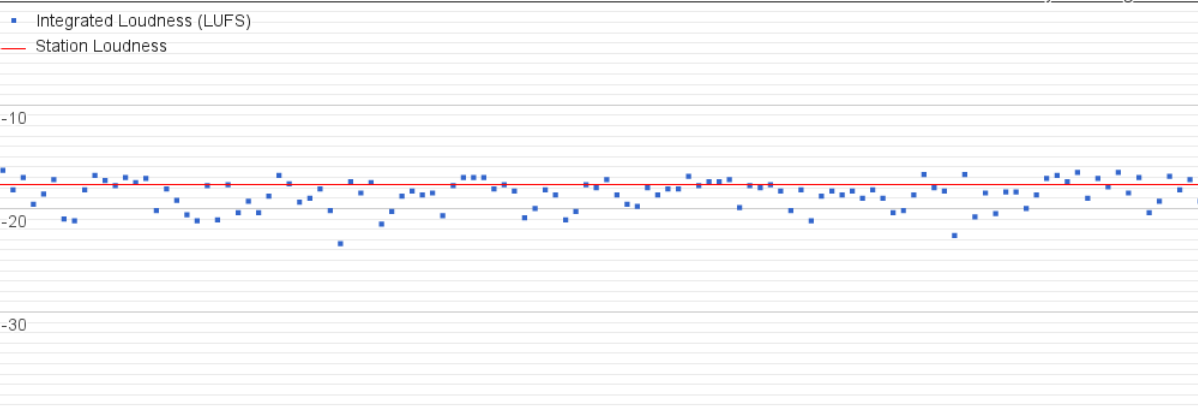
Mostwedding - Loudness Distribution

Adjacent Programmes



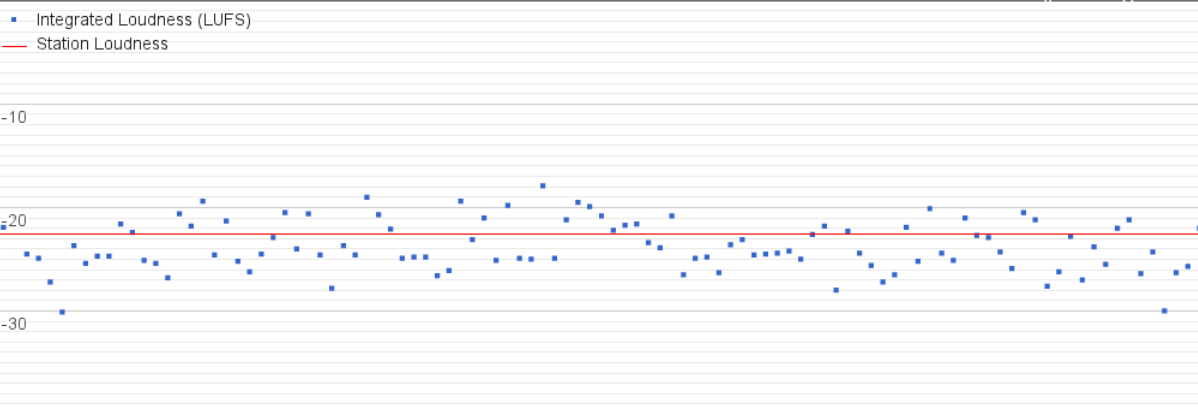
Jacaranda - Loudness Distribution

Adjacent Programmes



RSG - Loudness Distribution

Adjacent Programmes



5FM - Programme Loudness Distribution

Adjacent Programmes

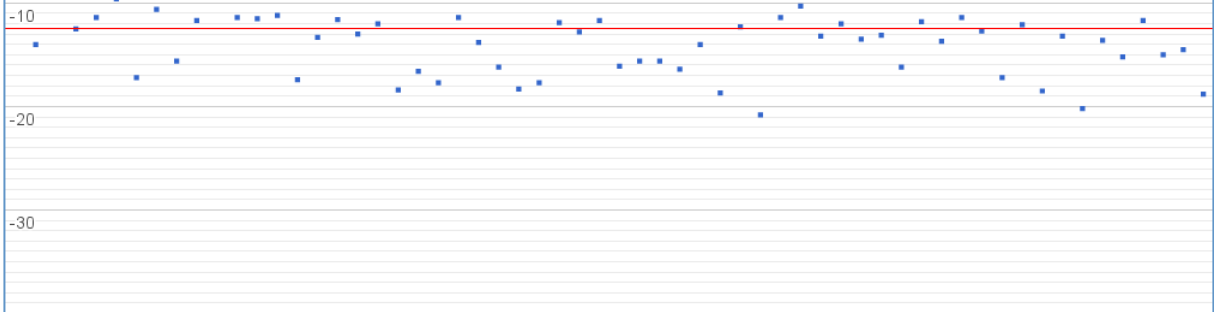
- Integrated Loudness (LUFS)
- Station Loudness



Ikwekwezi FM - Loudness Distribution

Adjacent Programmes

- Integrated Loudness (LUFS)
- Station Loudness



Radio 2000 - Loudness Distribution

- Integrated Loudness (LUFS)



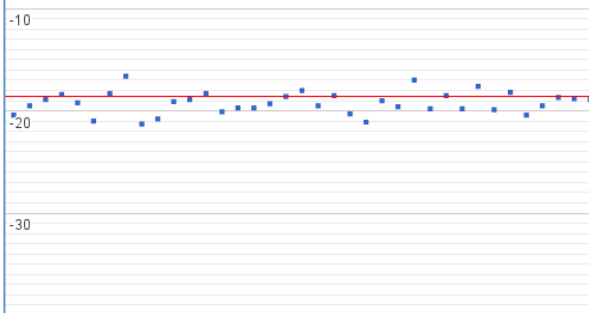
Power FM - Loudness Distribution

- Integrated Loudness (LUFS)



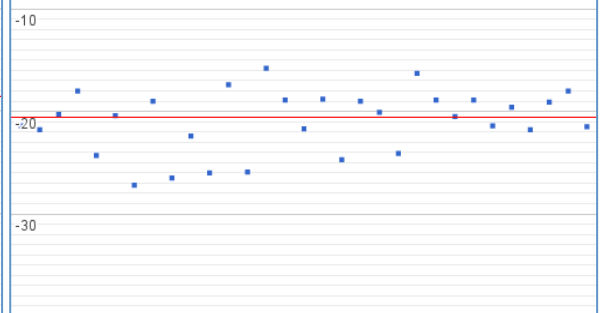
Classic FM - Loudness Distribution

- Integrated Loudness (LUFS)



SAfm - Loudness Distribution

- Integrated Loudness (LUFS)



Summary

An empirical investigation of loudness fluctuations in South African broadcast audio

By

Jozua Johannes Georg Loots

Supervisor: Mr Murray de Villiers

Department: Music

Degree for which this mini-dissertation is presented: Magister Musicae in Music Technology

The aim of the study was to investigate whether South African free to air television and radio broadcast audio contains loudness fluctuations that fall outside previously determined limits of listener comfort. This is a relevant aim as consumers often complain about loudness fluctuations in broadcast audio (e.g. “why are the commercials so loud?”). Loudness is an inherently subjective phenomenon that is not only subject to differences in human perception from day to day, but also more specifically by the frequency content, localisation, spatialisation and duration of the audio stimuli. Traditional audio level meters only measure the audio signal or digital samples, and do not take any of these psychoacoustic phenomena into consideration. Broadcast audio has traditionally been regulated by specifying the permitted maximum level (PML) of the audio to avoid overloading the transmitter or over-modulation the broadcast signals. While this is necessary to keep the transmission inside the technical dynamic range of the medium, it does not correspond to the perceived loudness of these signals. With the addition of powerful dynamic range processing techniques, content producers and broadcasters were now able to raise the average level (and correspondingly the perceived loudness) without affecting the permitted maximum level or the peak level of the signal. Broadcasts were still compliant, but subjectively louder. As this process has not been done uniformly across various stations, and various types of audio, fluctuations occur both between stations, and between different segments on the same station. These fluctuations are the cause of listener complaints.

There has been a move in international regulators and broadcasters to make a paradigm shift from peak normalisation to loudness normalisation of broadcast audio content. Limited, to no

adoption of this new paradigm is evident in South African broadcasting. This study provides baseline data of the *status quo* of South African free to air broadcast audio to investigate whether it contains problematic fluctuations, and therefore whether a move from peak to loudness normalisation could possibly have a positive effect.

The study found that generally radio broadcasts suffered from greater and more problematic loudness fluctuations compared to television. Televisions broadcasts differed enough from station to station to cause inter-station loudness fluctuations outside previously determined limits for listener comfort, but not intra-station fluctuations. SABC 2 was found to be the loudest and SABC 3 the softest at this particular testing high-site. Radio broadcasts contained a large proportion of inter-station fluctuations, and while it varied considerably from station to station, each station contains some proportion of intra-station fluctuations. Advertisements were found not to be the loudness programme segment type. Instead, segments type loudness was generally as follows (from softest to the loudest): talking, advertisements, links (interstitials) and music. Ikwewezi FM was found to be the loudest station by far, with Lotus FM the softest, with a very wide difference of 13.7LU between their integrated station loudness values. The vast majority of broadcast audio was found to have a loudness range appropriate to the intended platform, but perhaps limited to the range appropriate to the lowest common denominator platform, achieved by producing signals with high to extreme levels of dynamic compression and peak limiting.

The study also provided much greater levels of detail of the nature and extent of all loudness fluctuations, especially for radio loudness data. Additionally, ‘zap testing’ methodology was tested to simulate real-life inter-station fluctuation scenarios, and also found to be an efficient method for extrapolating overall station loudness for a larger set of stations.

Finally, the study recommends the paradigm shift from peak to loudness normalisation for all audio content producers and distributors in the country, and suggests the EBU R128 recommendation as the most viable starting point.

Keywords

Audio levels, normalisation, loudness, broadcast audio, loudness advocacy, permitted maximum level, loudness fluctuations, listening comfort zone, EBU R128, ITU-R BS.1770, integrated loudness, radio, television