
Mobile Handset Performance Evaluation Using Voice Quality Measurement White Paper

Metrico Wireless
7340 Executive Way
Suite U
Frederick, MD 21704



Contents

1	Abstract.....	3
2	Introduction	3
3	Measurement Methods.....	3
3.1	Voice Quality Measurement.....	3
3.2	Comparison to a Reference Mobile Handset Performance	3
3.3	Electrical Coupling.....	4
3.4	Testing Across Radio Environments and Infrastructure Types.....	4
4	Performance Evaluation	4
4.1	Comparison of Voice Quality Distribution Between the Reference and Subject Handset	5
4.2	Correlation of Voice Quality to Radio Environment	5

1 Abstract

This paper outlines methods for evaluating wireless handset performance through the analysis of voice quality measurements. In particular, the paper considers voice quality measurements collected from live-network tests that are designed to capture the user-experience. Discussion on methods for voice quality measurement, field-based test procedures and performance indicators is presented.

2 Introduction

Evaluating a mobile handset’s performance using voice quality measurement provides insightful understanding of over-all handset performance and user-experience. When performed in a live network setting, voice quality measurement offers the advantage of exercising both the radio subsystem and baseband processing subsystem for a mobile handset. This contrasts with testing that measures bit error rate and/or frame error rate and focuses on exercising only the radio subsystem. The baseband processing subsystem performance, including the voice codec and related voice processing, has been found to vary across mobile handset models; resulting in a varied user experience.

3 Measurement Methods

3.1 Voice Quality Measurement

Voice quality is effectively measured using standardized models that predict quality scores traditionally determined from subjective tests with human listeners. PESQ, an ITU-T recommendation P.862 standardized model for predictive speech quality, utilizes speech signals that are subjected to the communications link under test and compares the degraded speech to the original speech. Speech quality scores are provided on a 5-point scale (MOS-like listening quality score).

Score	Quality of Speech
5	Excellent
4	Good
3	Fair
2	Poor
1	Bad

3.2 Comparison to a Reference Mobile Handset Performance

When performed in a live network setting, evaluation of performance is enhanced when a subject handset is compared to a reference handset in a side-by-side test. Without a reference handset, temporary events that impact radio conditions at the time of test, may adversely affect a subject handset’s performance. Using a reference handset therefore assists in isolating “network-related” performance from “handset-related” performance.

Additionally, when evaluating empirical voice quality measurements, analysis is made more relevant when considered in relation to a known good-performing handset.

3.3 Electrical Coupling

When measuring voice quality performance in a field-based test scenario, electrical coupling is superior to acoustic coupling. Electrical coupling eliminates the potential for introduction of unwanted sounds affecting the voice quality scores. For example, if engine sound were introduced through the acoustic coupling apparatus, then the voice quality scores would be correlated to engine speed; an independent variable that is not related to handset performance.

3.4 Testing Across Radio Environments and Infrastructure Types

In order to best characterize a mobile handset's voice quality performance, it is desirable to exercise the handset over a plurality of radio conditions and infrastructure types. In this way, the voice quality measurements are sampled in a manner that is representative of over-all user experience within the live network.

4 Performance Evaluation

Voice quality measurements collected through empirical study can be evaluated using common statistical inference methods. Statistical comparisons of the average voice quality and standard deviation of voice quality from a particular set of samples may be considered for statistical inference conclusions.

In addition, there are a number of insightful performance indicators that in summation provide evidence for over-all performance conclusions. Two of many such performance indicators are outlined.

4.1 Comparison of Voice Quality Distribution between the Reference and Subject Handset

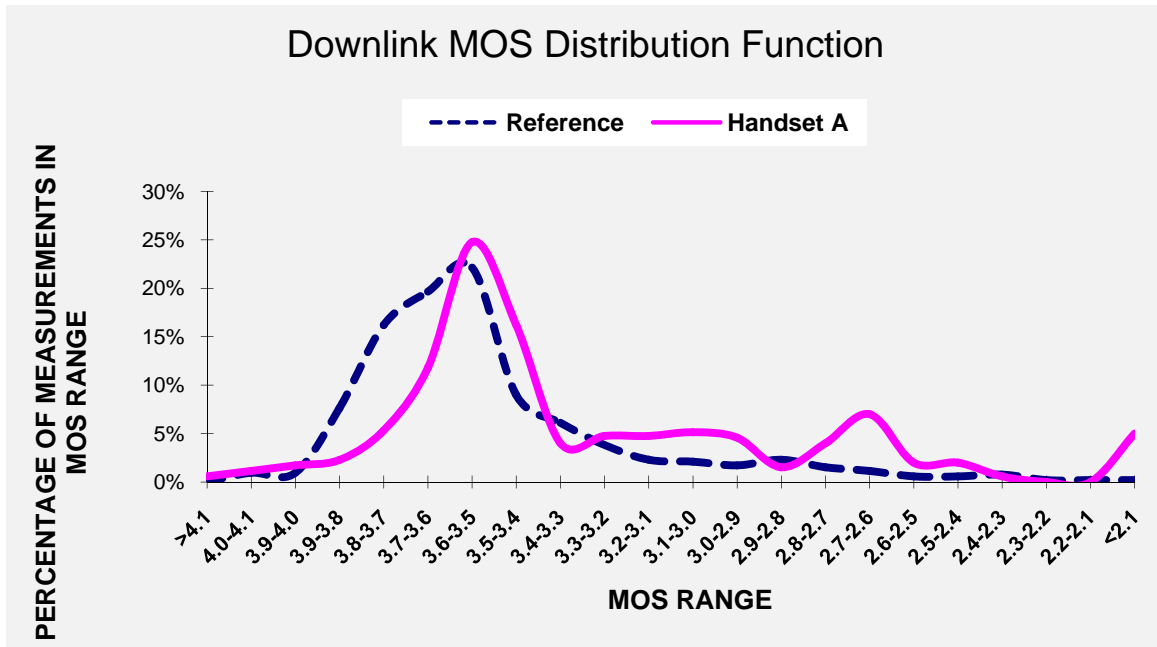


Figure 1- Handset A is likely to provide a more varied voice quality experience for the consumer than the Reference handset

4.2 Correlation of Voice Quality to Radio Environment

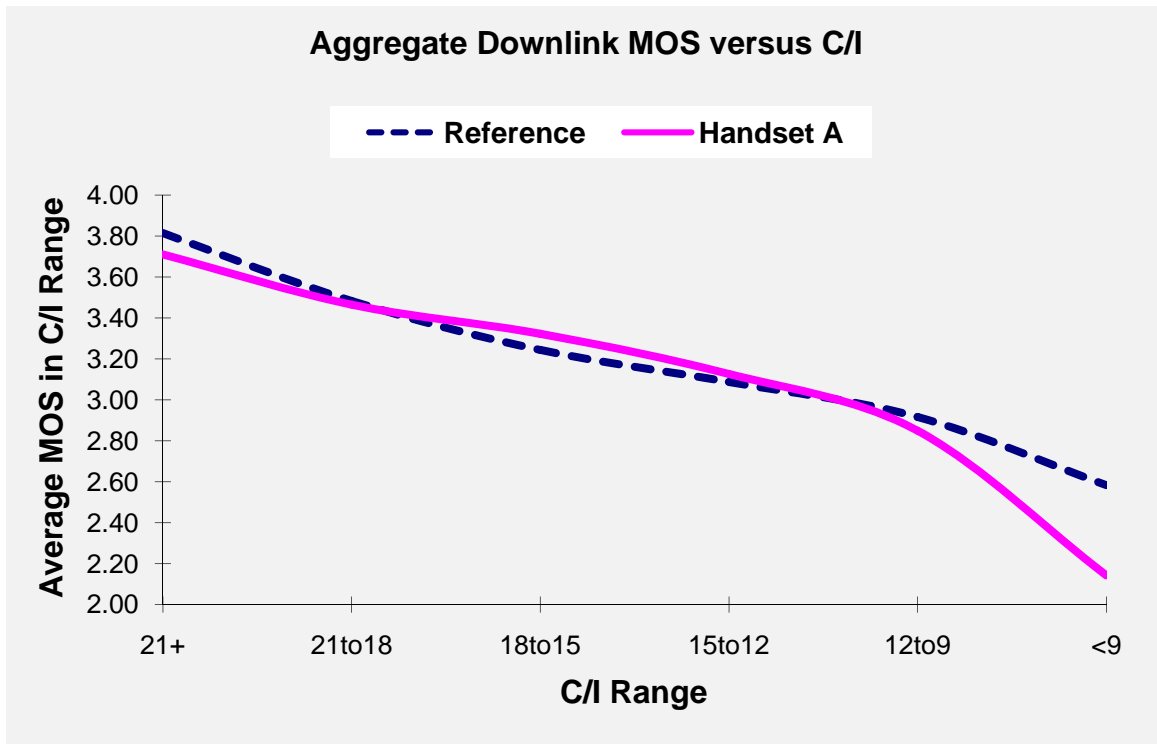


Figure 2 - Handset A is likely to provide a degraded voice quality experience in high interference environments compared to the Reference handset