ARIB STD-T64-C.S0056-A v1.0

Electro-Acoustic Recommended Minimum Performance Specification for cdma2000 mobile stations

Original Specification

This standard, ARIB STD-T64-C.S0056-A v1.0, was prepared by 3GPP2-WG of Association of Radio Industries and Businesses (ARIB) based upon the 3GPP2 specification, C.S0056-A v1.0.

Modification to the original specification

None.

Notes

None.

3GPP2 C.S0056-A v1.0

March 2013

3RD GENERATION PARTNERSHIP PROJECT 2 "3GPP2"

Electro-Acoustic Recommended Minimum Performance Specification for cdma2000 mobile stations

Revision History

Revision	Description of Changes	Date
Rev 0 v1.0	Initial Publication	July 2005
Rev A v1.0	Publication version of Revision A, version 1.0	March 2013

CONTENTS

TABLES

FIGURES

FOREWORD

This foreword is not part of this document.

This document was prepared by the Third Generation Partnership Project 2 (3GPP2).

.

1 INTRODUCTION

1.1 Scope

This document details definitions, methods of measurement, and minimum electro-acoustic performance characteristics for Code Division Multiple Access (CDMA) mobile stations. These standards are intended to ensure a level of electro-acoustic performance approaching that defined by the ITU-T for PSTN circuits. These electro-acoustic requirements are applicable to all Speech Service Options supported by the mobile station. These requirements are applicable to mobile station handsets that send and receive sound propagating exclusively through air. Requirements for handsets that send or receive sound by other means (e.g. tissue conduction or bone conduction), loudspeaker, headset, and hands-free configurations are not addressed in the current document and are for further study. Test methods are recommended in this document; however, methods other than those recommended may suffice for the same purpose. The electro-acoustic requirements in this document are not applicable to SO77 operation.

1.2 Document Conventions

1.2.1 Requirements Language

"Shall" and "shall not" identify requirements to be followed strictly to conform to this document and from which no deviation is permitted. "Should" and "should not" indicate that one of several possibilities is recommended as particularly suitable, without mentioning or excluding others, that a certain course of action is preferred but not necessarily required, or that (in the negative form) a certain possibility or course of action is discouraged but not prohibited. "May" and "need not" indicate a course of action permissible within the limits of the document. "Can" and "cannot" are used for statements of possibility and capability, whether material, physical or causal.

1.2.2 Other Conventions

1. "Base station" refers to the functions performed on the land side, which are typically distributed among a cell, a sector of a cell, and a mobile communications switching center.

2. The following operators define mathematical operations:

x indicates multiplication.

/ indicates division.

+ indicates addition.

- indicates subtraction.

* indicates complex conjugation.

$\in$ indicates a member of the set.

$\lfloor x \rfloor$ indicates the largest integer less than or equal to x: $\lfloor 1.1 \rfloor = 1$, $\lfloor 1.0 \rfloor = 1$.

|x| indicates the absolute value of x: |-17|=17, |17|=17.

1.3 References

1.3.1 Normative References

The following standards contain provisions which, through reference in this text, constitute provisions of this Standard. At the time of publication, the editions indicated were valid. All standards are subject to revision, and parties to agreements based on this Standard are encouraged to investigate the possibility of applying the most recent editions of the standards indicated below. ANSI and TIA maintain registers of currently valid national standards published by them.

1. ANSI S1.4-1983 (R2006), *American National Standard Specification for Sound Level Meters*.
2. ANSI S1.4A-1983 (R2006), *Amendment 1 - American National Standard Specification for Sound Level Meters*.
3. IEEE SA 269-2010, *Standard Method for Measuring Transmission Performance of Analog and Digital Telephone Set, Handsets and Headset*, August 2010.
4. ITU-T Recommendation G.122, *Influence of National Systems on Stability and Talker Echo in International Connections*, 1993.
5. TIA/EIA-690, *Recommended Minimum Standards for 800 MHz Subscriber Units*, November 2000.
6. 3GPP2 C.S0014-E v1.0, *Enhanced Variable Rate Codec, Speech Service Options 3, 68, 70, 73, and 77 for Wideband Spread Spectrum Digital Systems*, December 2011.
7. 3GPP2 C.S0011-E v1.0, *Recommended Minimum Performance Standards for cdma2000 Spread Spectrum Mobile Stations*, April 2012.
13. ITU-T Recommendation P.501, *Test signals for use in telephonometry*, January 2012.
15. ITU-T Recommendation P.501 Amd.1, *Test signals for use in telephonometry*, July 2012.
16. 3GPP TS 26.131 v11.0.0, *Terminal Acoustic Characteristics for Telephony; Requirements*, September 2012.
17. 3GPP TS 26.132 v11.0.0, *Speech and video telephony terminal acoustic test specification*, September 2012.
18. ITU-T Recommendation P.57, *Artificial Ears*, December 2011.
19. ITU-T Recommendation P.58, *Head and torso simulator for telephonometry*, December 2011.
20. ETSI TS 103 106, *Speech quality performance in the presence of background noise: Background noise transmission for mobile terminals-objective test methods,* August 2012.
21. ETSI EG 202 396 – 1 v1.2.4, *Speech quality performance in the presence of background noise; Part 1: Background noise simulation technique and background noise database,* February 2011.

1.3.2 Informative References

At the time of publication, the editions indicated were valid. All standards are subject to revision, and parties to agreements based on this Standard are encouraged to investigate the possibility of applying the most recent editions of the standards indicated below. ANSI and TIA maintain registers of currently valid national standards published by them.

8. IEEE Std 661-1979 (R2008), *IEEE Standard Method for Determining Objective Loudness Ratings of Telephone Connections*, September 2008.
9. ITU Recommendation P.79, *Calculation of loudness ratings for telephone sets*.
10. TIA/EIA-810-B, *Telecommunications - Telephone Terminal Equipment-Transmission Requirements for Narrowband*, November 2006.
11. TIA/EIA-579-A, *Telecommunications Telephone Terminal Equipment Transmission Requirements for Digital Wireline Telephones*, November 1998.
12. IEC 60950-1, *Information Technology Equipment – Safety –Part 1: General Requirements*.
13. Reserved.
14. TIA-470.110-C-1, *Telecommunications — Handset Acoustic Performance Requirements for Analog Terminal Equipment Connected to the Public Switched Telephone Network*, April 2007.
22. S4-120621, *Ext_ATS Permanent Document (EATS-3): Common subjective test framework for validation of P.835 test predictors (v.1.1),* Source: Editor (Qualcomm, Incorporated).
23. ITU-T P.1401, *Methods, metrics and procedures for statistical evaluation, qualification and comparison of objective quality prediction models*, July 2012.
24. S4-120841, *Ext_ATS Permanent document (EATS-6): Requirements for the objective P.835 predictor model performance and P.835 database collection,* Source: Editor (ORANGE; Qualcomm, Incorporated).
25. ITU-T T09-SG12-C-0305, *Application of P.863 SWB mode on conditions with noise reduction from mobile context, comparison with subjective test scores according to P.835,* Source: France Telecom Orange.
26. ITU-T P.863, *Perceptual Objective Listening Quality Assessment,* January 2011
27. ITU-T P.835, Subjective test methodology for evaluating speech communication systems that include noise suppression algorithm, November 2003

1.4 Terms

Base Station. A fixed station used for communicating with mobile stations. Depending upon the context, the term base station may refer to a cell, a sector within a cell, an MSC, or other part of the wireless system.

Codec. The combination of an encoder and decoder in series (encoder/decoder).

dBA. A-weighted sound pressure level expressed in decibels obtained by the use of a metering characteristic and the weighting A specified in [1] and [2].

dBm. A measure of power expressed in terms of its ratio (in dB) to one milliwatt.

dBm0. Power level expressed in dBm relative to an arbitrarily defined reference point called the zero transmission level point, or 0 TLP [3].

dBPa. A measure of sound pressure level expressed in terms of its ratio (in dB) to one Pascal, 20 log10 (Pressure/1 Pa). 1 Pa = 1Newton/m^2.

dBSPL. A measure of Sound Pressure Level expressed in terms of its ratio in (dB) to 20 µPa, 20 log10 (Pressure/20µPa). dBPa is preferred.

Drum Reference Point (DRP). Point located at the end of the ear canal, corresponding to the ear-drum position.

Decoder. A device for the translation of a signal from a digital representation into an analog format, for the purposes of this standard, a cdma2000®[1] compatible device.

Ear Reference Point (ERP). See section 3.20 of [3].

Encoder. A device for the coding of a signal into a digital representation, for the purposes of this standard, a cdma2000 compatible device.

Mobile Station. A station intended to be used while in motion or during halts at unspecified points. Mobile stations include portable units (e.g., hand-held personal units) and units installed in vehicles.

Mouth Reference Point (MRP). See section 3.30 of [3].

Nominal Volume Control Setting. Reference Receive Volume Control Setting, see section 3.38 of [3].

RLR. Receive Loudness Rating, a measure of receive audio sensitivity.

SLR. Send Loudness Rating, a measure of transmit audio sensitivity.

.

---

1 *cdma2000® is the trademark for the technical nomenclature for certain specifications and standards of the Organizational Partners (OPs) of 3GPP2. Geographically (and as of the date of publication), cdma2000® is a registered trademark of the Telecommunications Industry Association (TIA-USA) in the United States.*

2 ELECTRO-ACOUSTIC REQUIREMENTS

2.1 Receive Audio Interface

2.1.1 Receive Audio Frequency Response

2.1.1.1 Definition

The audio frequency response of the mobile station earpiece is defined as the variation of the ratio of the acoustic output of the earpiece to the signal input at a reference base station as a function of frequency. The measurement is conducted at DRP with diffuse-field correction.

2.1.1.2 Method of Measurement

The receive audio frequency response shall be measured according to [3], using a real speech test signal and level as described in 3.3.1.

2.1.1.3 Minimum Standard

The receive audio frequency response for narrowband and wideband operation shall lie within the floating masks defined by the respective mandatory upper and the mandatory lower boundary limits specified in Table 2.1-1. The points in the table shall be interpolated linearly on a log frequency scale to form piecewise linear upper and lower bounds. The measurement shall be conducted at DRP on a HATS compliant to [19] with 1/3rd octave bandwidth analysis. The diffuse-field correction of Table 2.1-2 shall be applied. The diffuse-field correction is per [19]. The receive frequency response masks shall apply at the nominal volume control setting, see 2.1.2.3.

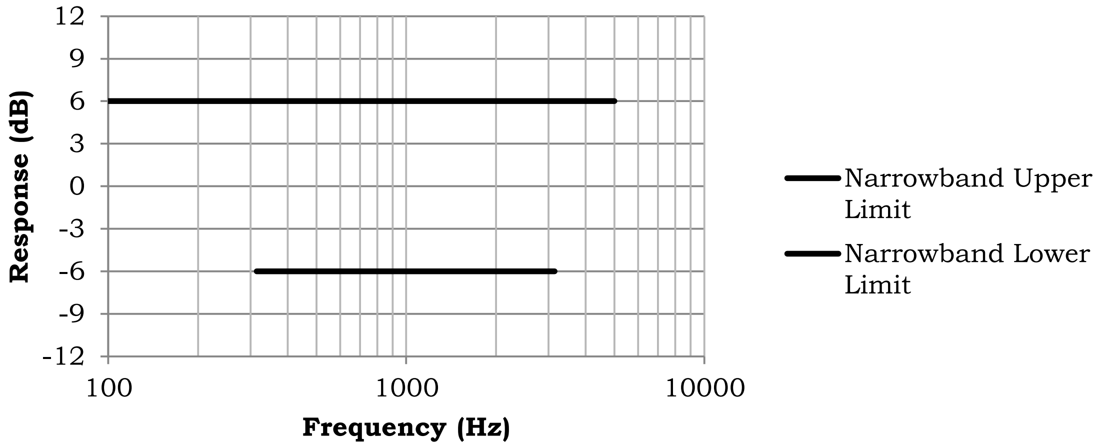
Table 2.1-1 Receive Frequency Response Mask

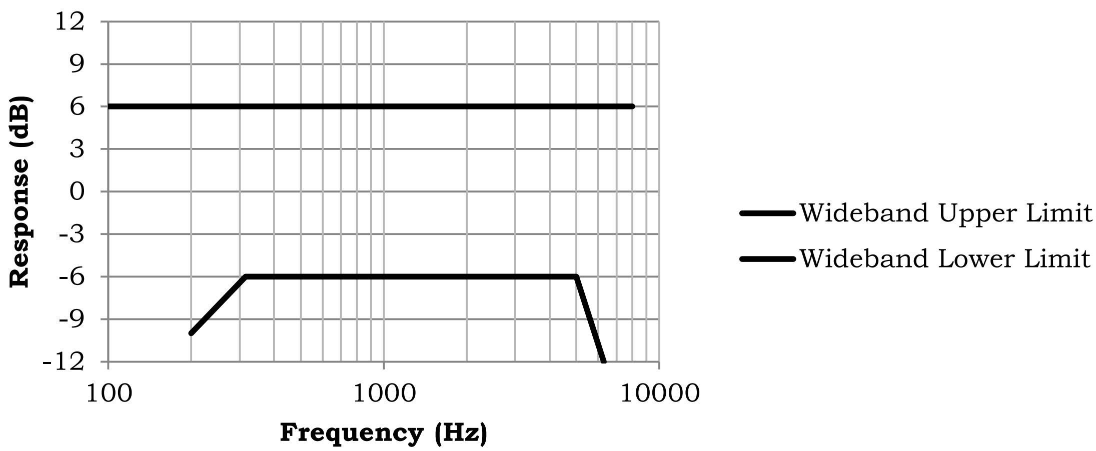
Frequency (Hz)	Mandatory Narrowband Upper Limit (dB)	Mandatory Narrowband Lower Limit (dB)	Mandatory Wideband Upper Limit (dB)	Mandatory Wideband Lower Limit (dB)
100	6		6	
200	6		6	-10
315	6	-6	6	-6
3150	6	-6	6	-6
4000	6		6	-6
5000	6		6	-6
6300			6	-12
8000			6	

Table 2.1-2 - Diffuse-field correction for receive frequency response measurements

Frequency (Hz)	HATS Diffuse Field correction (dB)
100	0
125	0
160	0
200	0
250	-0.5
315	-0.5
400	-1.0
500	-1.5
630	-2.0
800	-4.0
1000	-5.0
1250	-6.5
1600	-8.0
2000	-10.5
2500	-14.0
3150	-12.0
4000	-11.5
5000	-11.0
6300	-8.0
8000	-6.5

Figure 2.1-1 Receive Frequency Response Masks





2.1.2 Receive Audio Sensitivity

2.1.2.1 Definition

The receive audio sensitivity is the ratio of the acoustic output of the mobile station earpiece to the electrical input at a reference base station expressed as a Receive Loudness Rating (RLR). The reference point shall be ERP for this measurement.

2.1.2.2 Method of Measurement

The measurements shall be performed at both nominal and maximum volume settings. RLR is measured according to the procedure given in [3] using a real speech test signal and level as described in 3.3.1.

2.1.2.3 Minimum Standard

RLR shall be between -1 dB and +5 dB, and the nominal target value should be 2 dB. Where a user controlled receive volume control is provided, the nominal volume control setting is the setting that results in an RLR value closest to the nominal RLR. In cases where consecutive volume control settings result in RLR values that are equally far from the nominal RLR, the louder setting shall be selected. With the receive volume control set to maximum, the RLR shall not be less than (louder than) -13 dB. With the receive volume control set to minimum the RLR shall not be higher than (quieter than) 18dB. These requirements apply to both narrowband and wideband operation.

2.1.3 Idle channel noise in receiving direction

2.1.3.1 Definition

The idle channel noise in the receiving direction is the A-weighted sound pressure level generated by the mobile station (measured at DRP with diffuse-field correction) when no signal is applied at the input of the reference base station.

2.1.3.2 Method of Measurement

The Idle channel noise in receiving direction shall be measured according to [17]. If a mobile station volume control is provided, it shall be adjusted to its maximum setting.

2.1.3.3 Minimum Standard

The idle channel noise in receiving direction shall be ≤ -54 dBPa(A). These requirements apply to both narrowband and wideband modes of operation.

2.2 Send Audio Interface

2.2.1 Send Audio Frequency Response

2.2.1.1 Definition

The audio frequency response of the mobile station transmitter is defined as the variation of the ratio of the output from a reference base station to the acoustic input at the MRP as a function of frequency.

2.2.1.2 Method of Measurement

The send audio frequency response shall be measured according to [3], using a real speech test signal and level as described in 3.3.1 and using the mouth simulator as specified in [3].

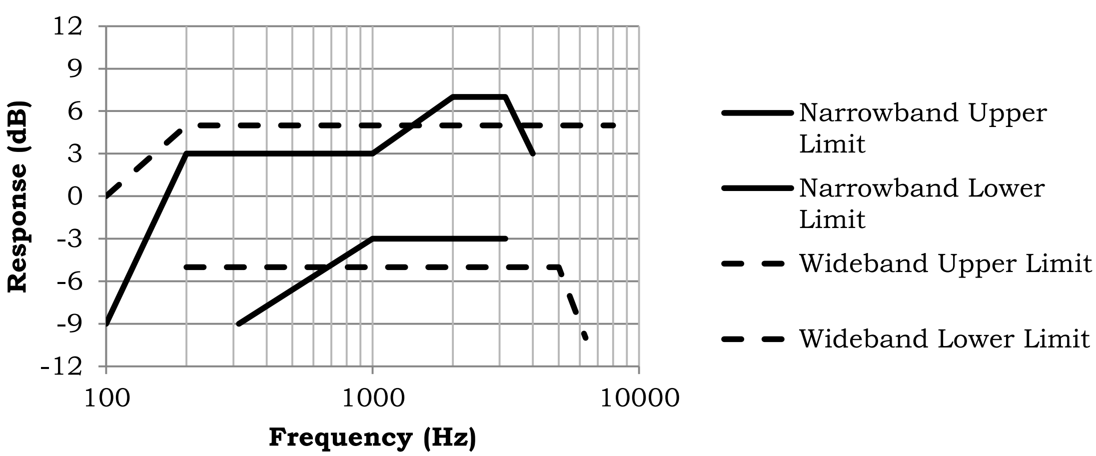
2.2.1.3 Minimum Standard

The send audio frequency response for narrowband and wideband operation shall lie within the floating masks defined by the respective mandatory upper and the mandatory lower boundary limits specified in Table 2.2-1. The points in the table shall be interpolated linearly on a log frequency scale to form piecewise linear upper and lower bounds. The measurement shall be conducted using 1/3rd octave bandwidth analysis.

Table 2.2-1 - Send Frequency Response Masks

Frequency (Hz)	Mandatory Narrowband Upper Limit (dB)	Mandatory Narrowband Lower Limit (dB)	Mandatory Wideband Upper Limit (dB)	Mandatory Wideband Lower Limit (dB)
100	-9		0	
200	3		5	-5
315	3	-9	5	-5
1000	3	-3	5	-5
2000	7	-3	5	-5
3150	7	-3	5	-5
4000	3		5	-5
5000			5	-5
6300			5	-10
8000			5	

Figure 2.2-1 Send Frequency Response Mask



2.2.2 Send Audio Sensitivity

2.2.2.1 Definition

The send audio sensitivity is the ratio of the acoustic sound pressure input at the MRP to the resulting output of a reference base station expressed as a Send Loudness Rating (SLR).

2.2.2.2 Method of Measurement

SLR is measured according to the procedure given in [3] using a real speech test signal and level as described in 3.3.1 and using the mouth simulator as specified in [3].

2.2.2.3 Minimum Standard

The SLR shall fall within the range +11 to +5 dB, and the nominal target value should be +8 dB. These requirements apply to both narrowband and wideband operation.

2.2.3 Weighted Terminal Coupling Loss

2.2.3.1 Definition

The weighted terminal coupling loss (TCL_W) is the ratio of the rms spectrum of the electrical input signal to the rms spectrum of the output signal of a reference base station simulator in conjunction with the mobile station under test. The echo signal results from acoustic coupling between the mobile station earpiece and the mobile station microphone.

2.2.3.2 Method of Measurement

The echo frequency response in a mobile station is measured in accordance with [3] using a speech-like test signal as described in 3.3.2.

For narrowband operation, the weighted terminal coupling loss is calculated from the echo frequency response according to [4] Annex B, Clause B.4.

For wideband operation, the weighted terminal coupling loss is calculated from the echo frequency response according to [4] Annex B, Clause B.4, but using the frequency range of 300 Hz to 6700 Hz (instead of 300 Hz to 3400 Hz).

2.2.3.3 Minimum Standard

With the receive volume control set to maximum, the weighted terminal coupling loss shall be at least 45 dB during single talk conditions. These requirements apply to both narrowband and wideband operation.

2.2.4 Idle channel noise in sending direction

2.2.4.1 Definition

The idle channel noise in the sending direction for narrowband operation is the psophometrically weighted electrical noise level generated by the mobile station (measured at the output of the reference base station) in silent conditions, i.e. when no acoustic signal is applied in the test room environment.

The idle channel noise in the sending direction for wideband operation is the A-weighted electrical noise level generated by the mobile station (measured at the output of the reference base station) in silent conditions, i.e. when no acoustic signal is applied in the test room environment.

2.2.4.2 Method of Measurement

The Idle channel noise in sending direction shall be measured according to the method described in [17].

2.2.4.3 Minimum Standard

The idle channel noise in sending direction shall be ≤ -64 dBm0p in narrowband operation.

The idle channel noise in sending direction shall be ≤ -64 dBm0(A) in wideband operation.

2.3 Loudness Contrast

2.3.1 Definition

Void

2.3.2 Method of Measurement

Void

2.3.3 Minimum Standard

Void

2.4 Wideband to Narrowband Loudness Contrast

2.4.1 Definition

Wideband to Narrowband Loudness Contrast is defined for both send and receive paths. Wideband to Narrowband Send Loudness Contrast is the difference between wideband SLR and narrowband SLR. Wideband to Narrowband Receive Loudness Contrast is the difference between wideband RLR and narrowband RLR. This test applies only to mobile stations that support both wideband and narrowband operation.

2.4.2 Method of Measurement

The measurements shall be performed at the nominal volume control setting determined in narrowband operation. Where a user controlled receive volume control is provided, the measurement shall also be performed at the maximum volume control setting.

Measure and record receive loudness rating per section 2.1.2 for wideband and narrowband operation.

Measure and record send loudness rating per section 2.2.2 for wideband and narrowband digital operation.

The volume control setting shall not be changed between the wideband and narrowband measurements.

2.4.3 Minimum Standard

The contrast in send and receive loudness ratings between wideband and narrowband operation for a given mobile station shall not exceed 3 dB.

2.5 Talker Sidetone characteristics

2.5.1 Talker Sidetone Masking Rating (STMR)

2.5.1.1 Definition

Talker Sidetone is the direction of speech transmission from mouth to ear of the telephone user. The Talker Sidetone Masking Rating is a single-number value which corresponds to the perceived loudness loss of the talker sidetone connection, and takes into account the human sidetone signal as a masking threshold, as specified in ITU-T Recommendation P.79.

2.5.1.2 Method of Measurement

The Talker Sidetone Masking Rating shall be measured according to the methods described in [17]; except that the test signal for normative measurements shall be the single-talk real speech test signal described in 3.3.1. If a mobile station in narrowband operation has a wideband sidetone path, then STMR shall be measured using the wideband test method.

2.5.1.3 Minimum Standard

The talker sidetone masking rating (STMR) shall be ≥ 15 dB and should be ≤ 23 dB for the nominal setting of the volume control. For all other positions of the volume control, the STMR shall be ≥ 10 dB.

2.5.2 Talker Sidetone Delay

2.5.2.1 Definition

The Talker Sidetone Delay is the difference between acoustical and electrical sidetone path delay measured between the HATS MRP and DRP.

2.5.2.2 Method of Measurement

The Talker Sidetone Delay shall be measured according to [17].

2.5.2.3 Minimum Standard

The maximum Talker Sidetone Delay shall be ≤ 10 ms, measured in an echo-free setup.

The measured result is only applicable where the level of the electrical sidetone is sufficiently high to be measured. While the STMR value may indicate the presence of sidetone it should be ensured that this is not primarily due to the acoustical or mechanical sidetone path when interpreting sidetone delay results.

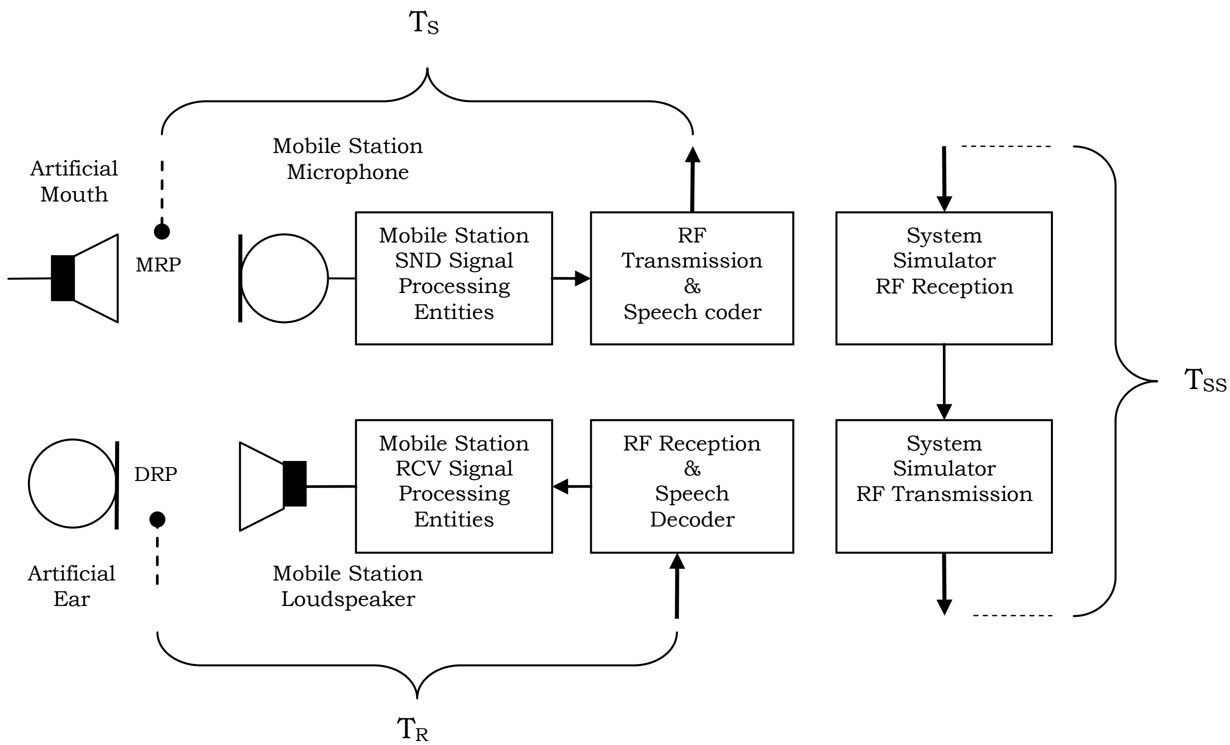
2.6 Mobile Station Delay

2.6.1 Definition

The mobile station delay is defined as the sum of the mobile station delay in sending and receiving directions, (T_S+T_R).

The mobile station delay shall include all entities from MRP to DRP (mouth-to-ear), but shall exclude the delays introduced by the test equipment and system simulator.

Figure 2.6-1 Mobile station delay measurement



The delay measured from MRP to DRP is (T_R + T_S + T_{SS}).

All system simulator delays, for the used network type, codec type and bitrate, (including radio access, speech codec, A/D and D/A conversions etc., added echo delay) are included in T_{SS}. The values used for testing (typical value considering variations due to interleaving etc.) as declared by the test equipment manufacturers shall be reported along with the measurement results.

2.6.2 Method of measurement

1. The setup of the mobile station is in correspondence to [3].

2. The system simulator is configured for “loopback” or “echo” operation. In “loopback” or “echo” operation, the signal in the sending direction is routed to the receiving direction by the system simulator.

3. A short word of real speech signal as specified in [13] section 7.3 is played back at

the mouth simulator and transmitted through the sending direction of the mobile station. The test signal level is -4,7 dBPa at the MRP.

4. The acoustic signal at MRP and the acoustic signal at DRP are acquired during the measurement.

5. The mouth-to-ear delay is determined by cross-correlation analysis between the measured signals at DRP and MRP. The measurement is corrected by subtracting the system simulator delay to obtain the mobile station delay.

6. The delay is measured in milliseconds and the maximum of the cross-correlation function is used for the determination.

2.6.3 Minimum Standard

The sum of the mobile station delays in sending and receiving directions (T_S + T_R) shall not exceed 220 ms and should not exceed 185 ms.

2.7 Speech Quality in sending in the presence of ambient noise

2.7.1 Definition

The speech quality in sending is evaluated according to [20]. This test method leads to three MOS-LQO quality numbers:

N-MOS-LQO: Transmission quality of the ambient noise

S-MOS-LQO: Transmission quality of the speech

G-MOS-LQO: Overall transmission quality

Although the model in [20] has been specifically trained with AMR speech coders, the model has shown to provide similar prediction performance with EVRC speech coders as described in ANNEX D:

2.7.2 Method of Measurement

The test arrangement is according to [21]. The measurement is conducted for 8 ambient noise conditions as described in Table 2.7-1. The measurements should be made in the same unique and dedicated call. The noise types shall be presented according to the order specified in Table 2.7-1.

Table 2.7-1 - Ambient noise types for measurement

Description	File name	Duration	Level	Type
Recording in pub	Pub_Noise_binaural_V2	30 s	L: 75.0 dB(A) R: 73.0 dB(A)	Binaural
Recording at pavement	Outside_Traffic_Road_binaural	30 s	L: 74.9 dB(A) R: 73.9 dB(A)	Binaural
Recording at pavement	Outside_Traffic_Crossroads_binaural	20 s	L: 69.1 dB(A) R: 69.6 dB(A)	Binaural
-	Clean (no noise)	-	L: <30 dB(A) R: <30 dB(A)	-
Recording at the drivers position	Fullsize_Car1_130kmh_binaural	30 s	L: 69.1 dB(A) R: 68.1 dB(A)	Binaural
Recording at sales counter	Cafeteria_Noise_binaural	30 s	L: 68.4 dB(A) R: 67.3 dB(A)	Binaural
Recording in a cafeteria	Mensa_binaural	22 s	L: 63.4 dB(A) R: 61.9 dB(A)	Binaural
Recording in business office	Work_Noise_Office_Callcenter_binaural	30 s	L: 56.6 dB(A) R: 57.8 dB(A)	Binaural

1. Before starting the measurements a proper conditioning sequence shall be used. The conditioning sequence shall be comprised of the four additional sentences 1-4 described in [20], applied to the beginning of the 16-sentence test sequence. The conditioning signal active speech level is - 1.7 dBPa at the MRP.

NOTE: The sequence of speech samples concatenated for the test signal, consisting of alternating talkers in the sending direction, reduces the overall test time but may represent an unrealistic behavior for certain voice enhancement technologies. Alternative concatenations can be used provided that a minimum of 4 different talkers is evaluated in the test.

2. The send speech signal consists of the 16 sentences of speech as described in [20]. The test signal active speech level is - 1.7 dBPa at the MRP. Three signals are required for the tests:

- The clean speech signal is used as the undisturbed reference (see [20], [21]).
- The speech plus undisturbed background noise signal is recorded at the terminal's primary microphone position using an omnidirectional measurement microphone with a linear frequency response between 50 Hz and 12 kHz.
- The send signal is recorded at the output of the reference base station.

3. N-MOS-LQO, S-MOS-LQO and G-MOS-LQO are calculated as described in [20] on a per sentence basis and averaged over all 16 sentences. The results shall be reported as average and standard deviation.

4. The measurement is repeated for each ambient noise condition described in Table 2.7-1.

5. The average of the results derived from all ambient noise types is calculated.

2.7.3 Minimum Standard

The mobile station shall comply with the following requirements:

S-MOS-LQO

- The average of S-MOS-LQO scores across all test conditions shall be ≥ 3.0.
- As a performance objective, the average of the S-MOS-LQO scores across all test conditions should be ≥ 3.5.

N-MOS-LQO

- The average of the N-MOS-LQO scores across all test conditions shall be ≥ 2.5.
- As a performance objective, the average of N-MOS-LQO scores across all test conditions should be ≥ 3.0.

G-MOS-LQO

- The average of the G-MOS-LQO scores across all test conditions shall be ≥ 2.5.
- As a performance objective, the average of G-MOS-LQO scores across all test conditions should be ≥ 3.0.

The requirements apply to both narrowband and wideband operation.

3 ELECTRO-ACOUSTIC STANDARD TEST CONDITIONS

3.1 Reference Base Station Simulator Requirements

3.1.1 Receive Level

A Digital Test Sequence (DTS) representing the reference base station simulator's codec equivalent of an analog sinusoidal signal whose rms value is 3.17 dB below the full-scale sinusoid capacity of the codec shall generate a level of 0dBm0 at the PSTN network [6], [3].

All signal levels will have a tolerance of ±0.5 dB as specified in [3].

All levels are specified in a 600 Ohm system [3].

3.1.2 Send Level

A sinusoidal signal with a level of 0dBm0 shall produce the digital test sequence (DTS) representing the codec equivalent of an analog sinusoidal signal whose rms value is 3.17 dB below the full-scale sinusoidal capacity of the codec [6], [3].

All signal levels will have a tolerance of ±0.5 dB as specified in [3].

All levels are specified in a 600 Ohm system [3].

3.1.3 Other Requirements

Other requirements for the reference base station simulator are defined in 6.4.3 of [7].

3.2 Ear Simulators

A Type 3.3 ear simulator (Anatomically shaped soft pinna according to [3]) shall be used for all tests.

3.3 Acoustic Test Signals

The test stimuli for the level, gain, frequency-response, and echo-loss measurements in this document are based on natural speech. The use of natural-speech based stimuli is predicated by the inherent nature of CDMA vocoders, which were designed for speech signals.

For narrowband operation, the test signal used shall be band limited between 100 Hz and 4 kHz with a bandpass filter providing a minimum of 24 dB/oct. filter roll-off, when feeding into the receiving direction.

For wideband operation, the test signal used shall be band limited between 100 Hz and 8 kHz with a bandpass filter providing a minimum of 24 dB/oct. filter roll-off, when feeding into the receiving direction.

3.3.1 Normal Test Signal

The default test stimulus for the level, gain, and frequency-response measurements is the British-English single talk sequence described in [13]. The overall active speech level of the composite signal shall be adjusted to a level of –16 dBm0 for receiving direction tests and -4.7dBPa at MRP for sending direction tests, unless otherwise specified.

3.3.2 Echo-Loss Test Signal

The default test stimulus for the echo-loss measurement is the compressed real speech signal described in clause 7.3.3 of [15]. The overall active-speech level of the composite signal shall be adjusted to a level of -10 dBm0.

3.4 Standard Test Conditions

All narrowband operation tests shall be performed with a call set up using RC3 and service option 3 or SO68 COP0 unless otherwise specified. All wideband operation tests shall be performed with a call set up using service option 73 COP0 if supported, otherwise SO70 COP0, unless otherwise specified. See 6.5.2 of [7] for comments that apply to all tests unless otherwise specified.

ANNEX A: [Informative] Loudness Rating Conversions

This annex is informative.

The loudness ratings defined in [8] may be approximated from those defined in [9], as described by [10], as follows:

SLR ([11]) = TOLR ([8]) + 57

RLR ([11]) = ROLR ([8]) - 51

Results obtained with the above conversions should be used for reference only. These conversions are based upon approximated frequency response curves as specified in [10]. Proper conversions may depend on actual measurements being made with each measurement standard where frequency responses deviate significantly from the norm.

This section has been referenced from Annex T3 of [3].

This page left intentionally blank

ANNEX B: [Informative] Maximum Acoustic Pressure Level

This annex is informative.

Maximum acoustic pressure requirements are defined in [12]. A test signal appropriate for use with the codecs employed by cdma2000 mobile stations is needed.

The recommended test stimulus for the short-duration measurement is PeakLevel.pcm (alternatively PeakLevel.wav). It is based on the 12burshp.pcm artificial signal that was provided as a potential candidate for general use as a stimulus in the measurements specified in this document. This 12-tone signal consists of the 65 lines of the P.501 [13] signal, broken into 12 vocals played in a random sequence. Each vocal consists of 5 or 6 lines that are harmonics of the fundamental. The vocals were additionally spectrally shaped to be more quasi-speech-like in nature.

While this signal proved less than satisfactory for use as a general stimulus in this MPS, a clipped version of this signal can be used as a stimulus for the specific peak acoustic level test. The signal is level adjusted in a hard limiting digitally-implemented amplifier such that the peak level of the composite signal is at the digital limits of the amplifier. The signal is then given a 25 dB gain. The resulting 25 dB clipped composite signal has 72.5% of its samples at the digital limits of the output of the amplifier.

When this stimulus is appropriately interfaced to the system, a measurable portion of the acoustic output of the mobile terminal will potentially be at the peak output level. An appropriately interfaced peak-sampling sound-pressure level meter can then be used to measure such peak acoustic transitions.

An appropriate test signal for the long-duration measurement is for further study.

This page left intentionally blank.

ANNEX C: [Informative] Ear Simulator

This annex is informative. See [18].

This page intentionally left blank.

ANNEX D: [Informative] Performance of the ETSI TS 103 106 model with SO68 COP0 test conditions

This Annex is informative.

D.1 Summary

[20] describes testing methodologies which can be used to objectively evaluate the performance of narrowband and wideband mobile terminals for speech communication in the presence of background noise.

In order to evaluate the applicability of [20] to CDMA2000 mobile stations employing SO68 COP0, two validation listening experiments according to the framework described in [22], but using CDMA2000 mobile stations employing SO68 COP0, were conducted. This contribution analyzes the performance of [20] as a predictor of the P.835 scores for these listening experiments.

The results indicate that [20] meets the performance targets established in [24] also for the CDMA2000 mobile stations employing SO68 COP0.

The prediction of wideband terminal performance and different speech coding rates are for further study.

The test conditions for the listening experiments were also processed with the Perceptual Objective Listening Quality Assessment (P.OLQA) algorithm described in [26].). It is found that, with a proper mapping function, the P.OLQA scores can also be used as a predictor of the [27] OVRL score (within the same performance constraints established in [24]).

D.2 Subjective data Normalization

During the Ext_ATS work item in 3GPP Rel.11, a framework [22] was developed to conduct listening experiments according to [27] for terminals. It is known [23] that, generally, listening experiment results cannot be directly compared to each other due to differences in the experiment context (e.g. range of degradations involved, number of test conditions) and other inter-lab aspects.

To illustrate this point, figures 1 to 3 present the OVRL (overall transmission quality), SIG (speech quality) and BAK (background noise intrusiveness) mean opinion score results for the reference conditions across 9 different [27] narrowband listening experiments conducted according to [22]. Each of the 9 experiments contain 12 reference conditions and 48 test conditions as described in [22] and all the experiments shared the same source speech material and processing for the reference conditions. The results shown include databases used for the training and validation of [20].

Figure D-1 OVRL scores for reference conditions of 9 different listening experiments according to 3GPP SA4 Ext_ATS work item subjective test framework

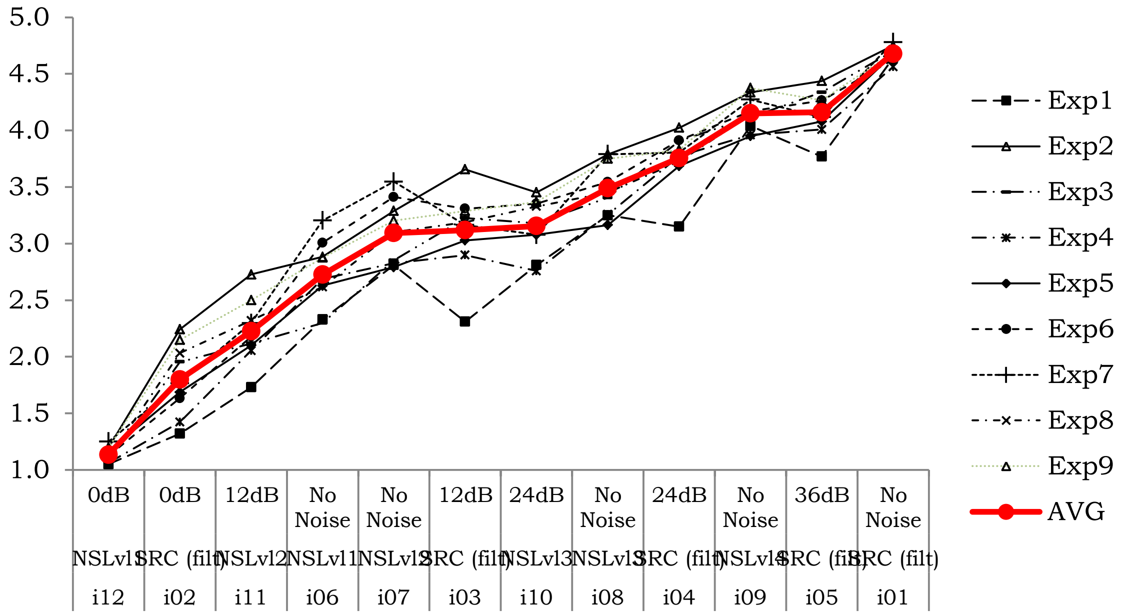


Figure D-2 SIG scores for reference conditions of 9 different listening experiments according to 3GPP SA4 Ext_ATS work item subjective test framework

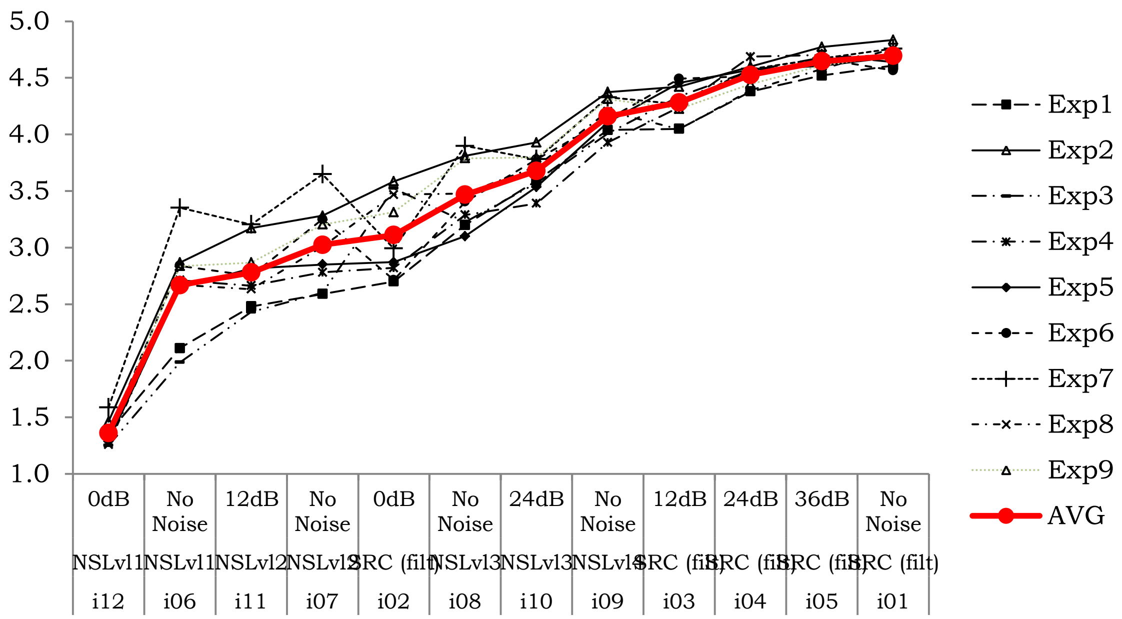
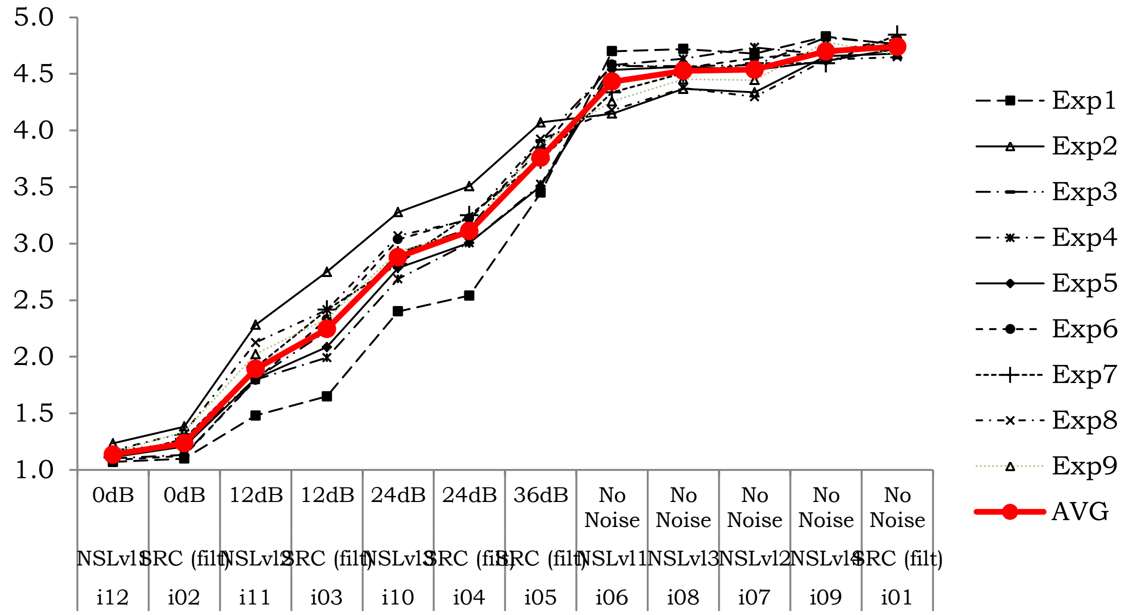


Figure D-3 BAK scores for reference conditions of 9 different listening experiments according to 3GPP SA4 Ext_ATS work item subjective test framework



It can be seen that rather large variations - sometimes in excess of 1.0 MOS - can exist from test to test. However, by maintaining the exact same recorded material for the set of reference conditions across multiple listening experiments (as it was done in [22]), it is possible to create a transformation polynomial function using the reference condition results of each individual experiment to the average of the reference conditions from multiple experiments. This “normalization” of the subjective data is meant to partially compensate for context based and inter-lab differences in the results and is applied to convert the raw [27] test condition scores into “normalized” [27]. A similar approach is described in [20], clause 5.

A second order polynomial is chosen for this normalization of the subjective data because the results between experiments generally match at the extremes of the scale (conditions with scores close to 1 and 5). In other words, the bias between the results of the tests gets compressed at the extremes of the scale and a 1st order normalization equation would not be appropriate to handle this aspect. The polynomial coefficients for each experiment are obtained through the least squares method.

Out of the 9 different narrowband experiments, experiment 1 is an [27] listening experiment containing 36 AMR 12.2kbps based test conditions and 12 SO68 COP0 test conditions, and experiment 3 is an [27] experiment containing 48 SO68 COP0 based test conditions. The remaining experiments contain AMR based only test conditions as they were used in the context of the 3GPP SA4 Ext_ATS work item. All the test conditions were recordings done with commercially available terminals.

The subjective data normalization functions for experiment 1 are given below.

$OVRL_{norm\_exp1} = -0.163+1.526*OVRL_{exp1}-0.107*OVRL_{exp1}^2$

$SIG_{norm\_exp1} = -0.429+1.557*SIG_{exp1}-0.099*SIG_{exp1}^2$

$BAK_{norm\_exp1}=-0.484+1.782*BAK_{exp1}-0.151*BAK_{exp1}^2$

The subjective data normalization functions for experiment 3 are found below:

$OVRL_{norm\_exp3} = -0.035+1.134*OVRL_{exp3}-0.033*OVRL_{exp3}^2$

$SIG_{norm\_exp3} = 0.327+1.065*SIG_{exp3}-0.033*SIG_{exp3}^2$

$BAK_{norm\_exp3}=0.025+0.995*BAK_{exp3}-0.004*BAK_{exp3}^2$

D.3 Objective data treatment

The objective scores are calculated according to the methods described in [20] clause 9, and the following 3rd order mapping equation is used.

$MOS\text{-}LQO_{norm} = a0+a1*MOS\text{-}LQO+a2*MOS\text{-}LQO^2+a3*MOS\text{-}LQO^3$

Mapping coefficients for Experiment 1 are per below:

	G-MOS-LQO	S-MOS-LQO	N-MOS-LQO
a0	17.23189	9.13415662	1.294398
a1	-12.543	5.59202531	0.590386
a2	3.558689	1.68050418	-0.00452
a3	-0.31035	0.14771248	0.006652

Mapping coefficients for Experiment 3 are per below:

	G-MOS-LQO	S-MOS-LQO	N-MOS-LQO
a0	15.81617	8.39831	11.75663
a1	-13.9994	-5.60634	-9.94218
a2	4.780718	1.890416	3.154512
a3	-0.50142	-0.18818	-0.28893

Scatter plots between the normalized subjective data and the predicted scores from [20] are shown in Figures 7 and 8.

Figure D-4 Scatter plot of normalized subjective scores and ETSI TS 103 106 objective predicted scores for Experiment 1

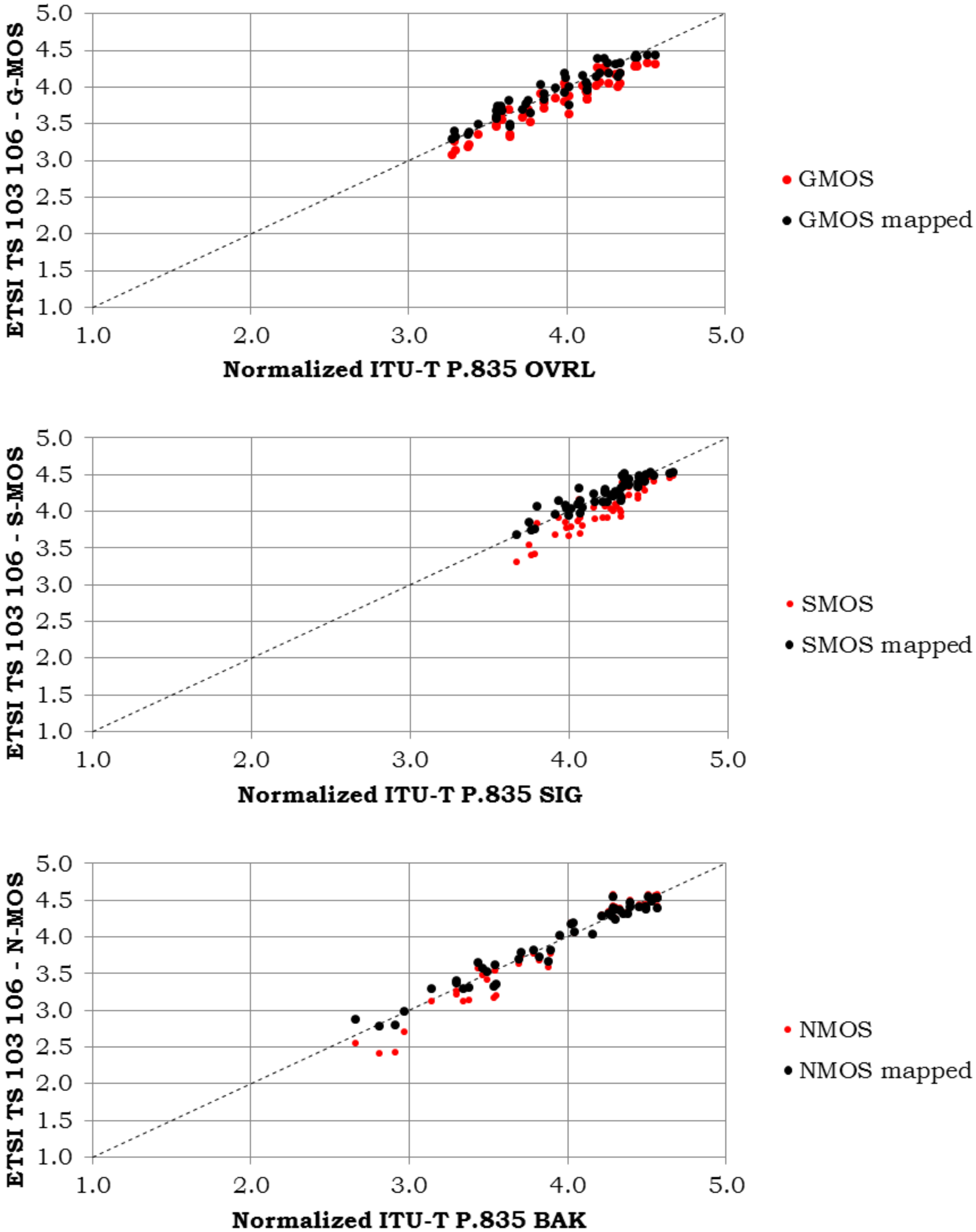
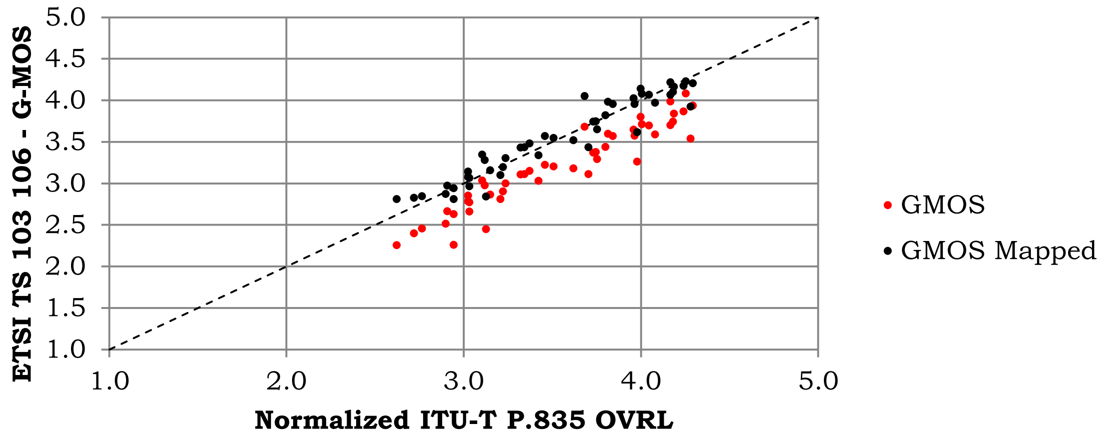
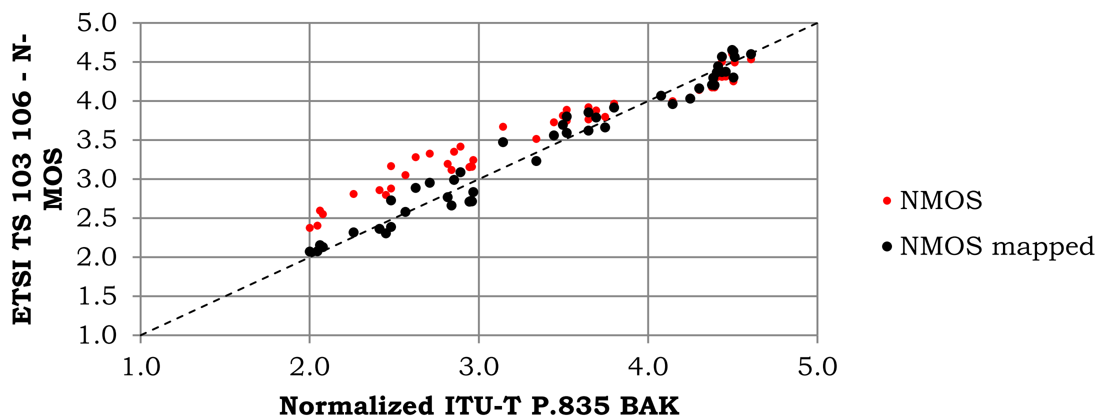
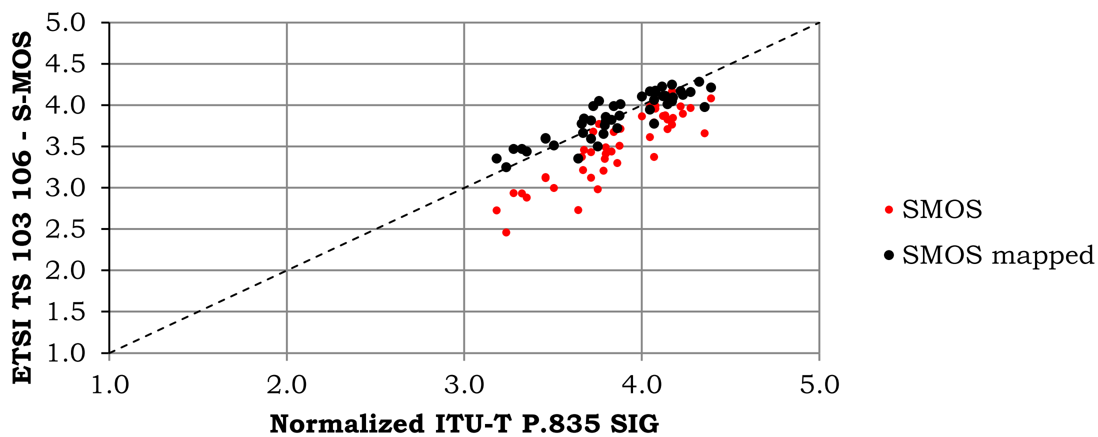


Figure D-5 Scatter plot of normalized subjective scores and ETSI TS 103 106 objective predicted scores for Experiment 3







D.4 Performance evaluation of objective predictor methodology

By following the objective predictor performance evaluation methodology described in [24], i.e. analysis of the RMSE after 3rd order mapping per experiment, the performance of the model with the two experiments containing SO68 COP0 test conditions is shown below:

Experiment 1

For experiment 1, the Pearson correlation coefficient between the [20] scores and the normalized [27] subjective scores is > 0.9 for all quality dimensions. For this particular experiment, the model seems to slightly under predict the S-MOS and G-MOS components. No specific over or under prediction for CDMA2000 mobile stations employing SO68 COP0 was observed in this particular experiment. The RMSE is < 0.35 prior to mapping of the objective scores and satisfy the criteria established in [24].

	OVRL	SIG	BAK
CORREL	0.95	0.91	0.98
RMSE	0.18	0.23	0.17
RMSE 3rd	**0.11**	**0.10**	**0.11**

Experiment 3

For experiment 3, the Pearson correlation coefficient of the [20] scores and the normalized [27] subjective scores is > 0.9 for all quality dimensions. For this particular experiment, the model seems to under predict the S-MOS and G-MOS components and over-predict the N-MOS. The RMSE is > 0.35 prior to mapping of the objective scores. After 3rd order mapping of the objective data however, the RMSE meets the requirements specified in EATS.

	OVRL	SIG	BAK
CORREL	0.95	0.97	0.94
RMSE	0.38	0.35	0.37
RMSE 3rd	**0.14**	**0.15**	**0.16**

While the performance prediction of the method in [20] comfortably exceeds the requirements defined in [24], it is not possible to conclude from this data that no over or under prediction of scores for CDMA2000 mobile stations employing SO68 COP0 exist. While the data from Experiment 1 (which includes both AMR and EVRC terminals) does not indicate any particular bias with regards to the choice of speech codec, the data in Experiment 3 is generally under predicted, even after the subjective test data normalization procedure. It is possible that either the specific subjective test conducted had lower scores than typical due to the context of the test or that an actual under prediction exists. Further studies are necessary to evaluate this aspect.

D.5 Performance of ITU-T P.863 as an objective predictor of OVRL quality

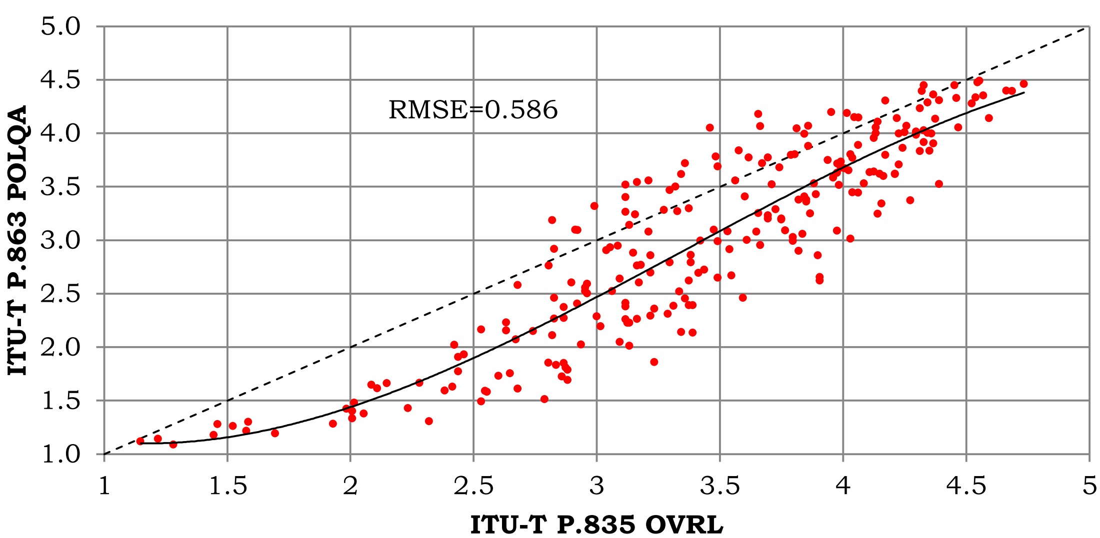
In [25], it has been reported that the [26] scores (P.OLQA) appeared to have good correlation to the OVRL and BAK [27] scores with the databases taken by the Source. The subjective scores results of 5 different experiments (240 test conditions) for narrowband terminals are used here to evaluate the performance of the [26] as a predictor of the [27] quality.

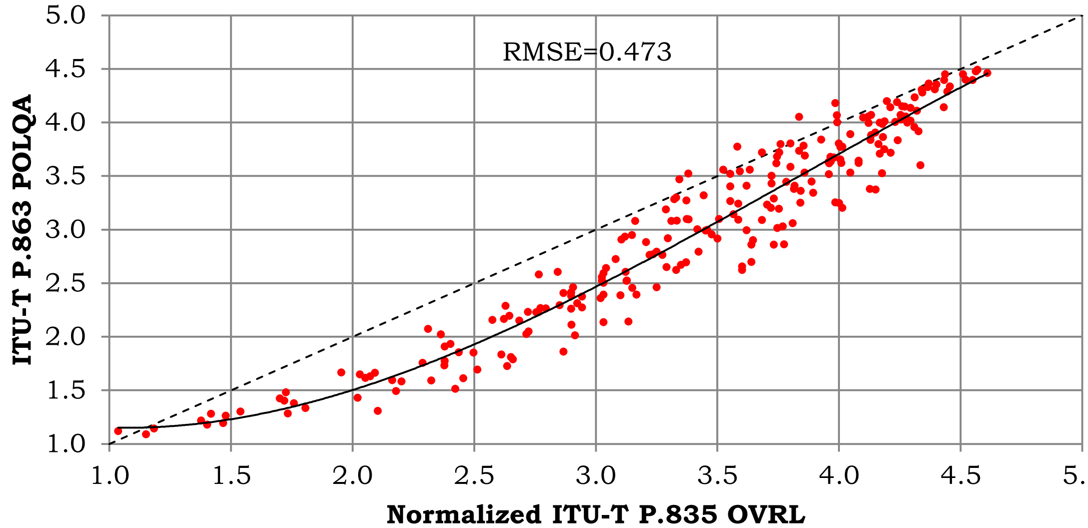
The following steps are taken:

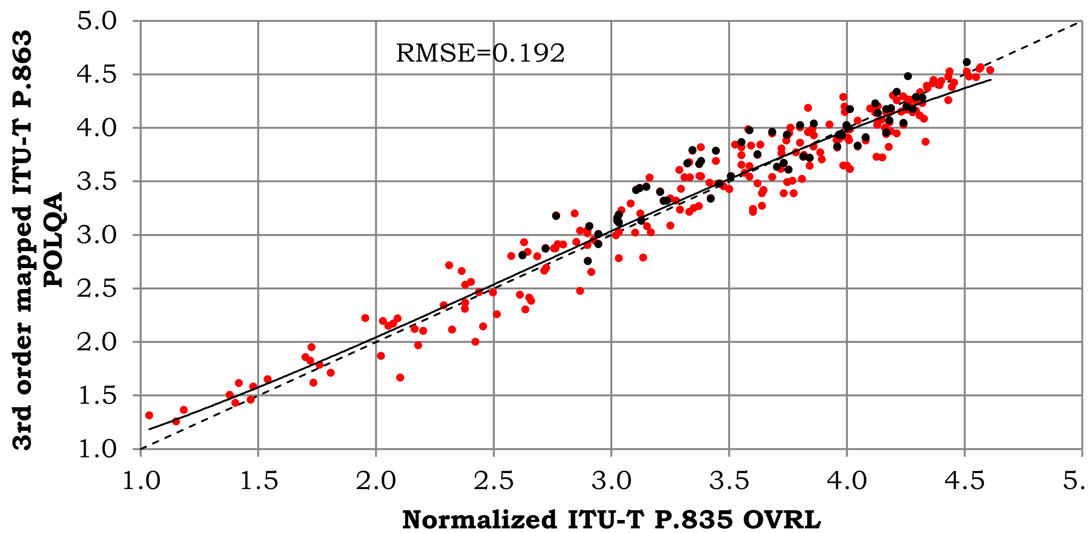
- Normalization of the subjective test data for each experiment as described in section D.2.
- Calculation of an [26] score per each 4s sample and then averaging across all samples per condition.
- A single monotonic 3rd order mapping derived from all 240 test conditions is applied to the [26] scores.
- The RMSE performance is evaluated according to the criteria in [24].

Results after each step are presented below:

Figure D-6 Scatter plots of P.OLQA scores x ITU-T P.835 OVRL scores







After the statistical data treatment described above, the following performance is obtained for the mapped P.OLQA as a predictor of the OVRL [27] scores:

Pearson Correlation Coefficient: 0.971

RMSE: 0.192

SO68 COP0 test conditions are presented as red points in the scatter plot.

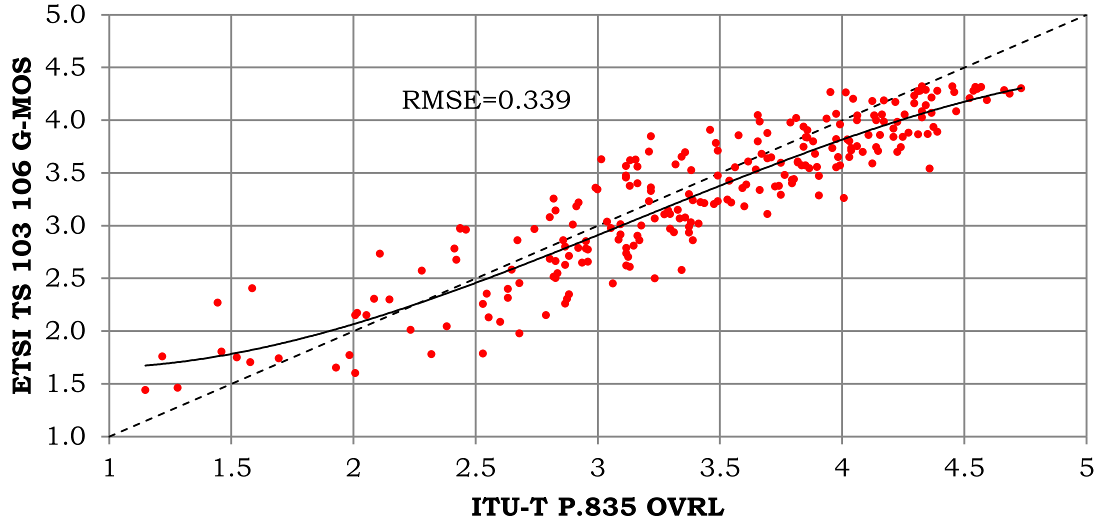
The equation used to map the raw P.OLQA scores to the OVRL subjective scores is given below:

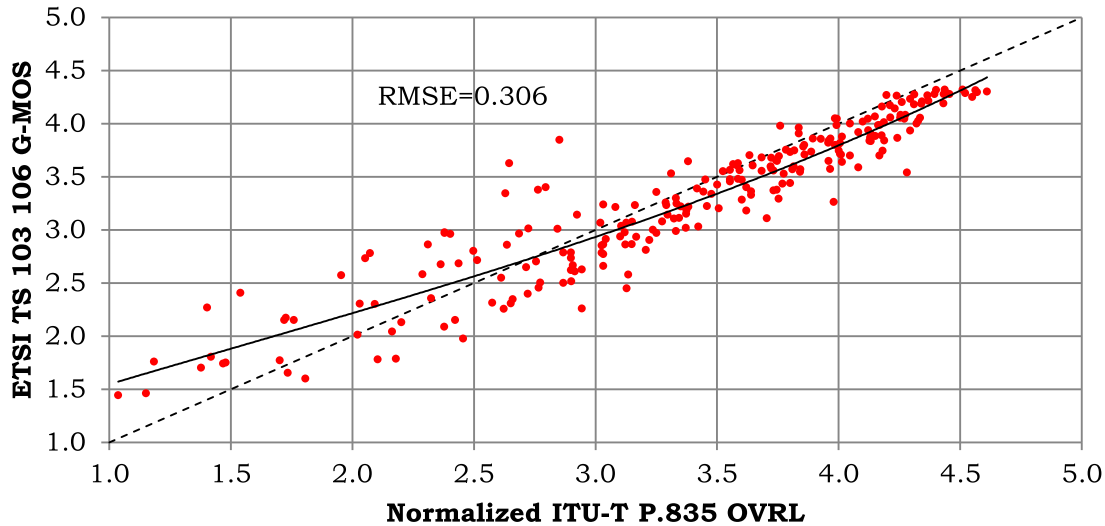
$G\text{-}MOS\text{-}LQO=-1.726+3.603*POLQA-0.888*POLQA^2+0.088*POLQA^3$

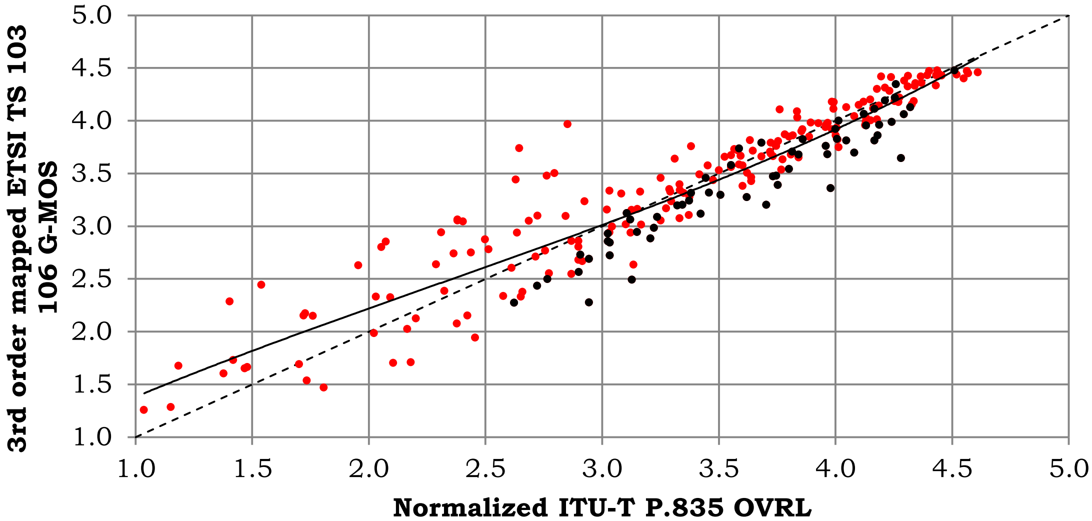
These results are considered in the context of the subjective test framework adopted in [22] only and should not be generalized for other scenarios, where the model in [26] has shown limitations (comparison of performance of different COPs, different speech coding technologies, etc.).

As a comparison point, the same procedure is applied to the model in [20]:

Figure D-7 Scatter plots of ETSI TS 103 106 scores x ITU-T P.835 OVRL scores







Pearson Correlation Coefficient: 0.935

RMSE: 0.287

Both methodologies (the mapped [26] scores presented here and [20]) match the performance requirements established in [24] for prediction of the OVRL scores.