

## ANEXO: METODOLOGÍA DE EVALUACIÓN MOS

### 1. Introducción

En el marco del proceso de actualización de la Norma Técnica de Calidad para la Prestación del Servicio Móvil Avanzado (SMA), ARCOTEL ha propuesto la modificación del indicador de Calidad de Conversación (MOS), incorporando el uso del algoritmo POLQA Full Band (ITU-T P.863) y estableciendo como valor objetivo que al menos el 90 % de las muestras obtengan un  $MOS \geq 2,4$ .

CONECCEL reconoce y respalda plenamente:

- La adopción de POLQA Full Band como metodología actualizada y alineada con estándares internacionales.
- El valor  $MOS \geq 2,4$  como umbral de inteligibilidad mínima, conforme lo establecido en la Recomendación ITU-T P.863.1 para códecs de referencia en Full Band.

No obstante, mi Representada recomienda que el enfoque planteado para la evaluación del cumplimiento corresponde una evaluación de la calidad de conversación mediante el Percentil 90 (P90) de la distribución de las muestras de MOS.

Este documento presenta los fundamentos técnicos, estadísticos y regulatorios que sustentan que la métrica más acertada con la QoE y con las propias recomendaciones internacionales citadas por ARCOTEL en el Informe de presentación del Proyecto Normativo, es el Percentil.

### 2. Naturaleza del indicador MOS - Métrica de Calidad de Experiencia (QoE)

La Recomendación ITU-T P.10 / G.100 define la Calidad de Experiencia (QoE) como el nivel de agrado o molestia percibido por el usuario final respecto del servicio recibido.

**6.209 quality of experience (QoE)**

The degree of delight or annoyance of the user of an application or service. [b-Qualinet2013]

En este contexto:

- MOS no mide parámetros físicos de red (QoS) como potencia, retardos, pérdida de paquetes, voltios, dBm, etc.
- MOS cuantifica la percepción subjetiva del usuario, consolidada estadísticamente. Es decir, mide qué tan buena o mala le pareció la comunicación al usuario.
- Cada opinión score representa una percepción individual.

Por tanto:

QoE es un concepto perceptual y MOS es el instrumento estadístico para medir la calidad de experiencia QoE.

Evaluar MOS como un KPI rígido de cumplimiento porcentual desnaturaliza su esencia, pues no considera la distribución real de la experiencia percibida y trata de forma idéntica eventos aislados y experiencias masivas.

### 3. Limitaciones técnicas del enfoque de “porcentaje de cumplimiento”

El indicador propuesto por ARCOTEL del tipo: “% de las muestras deben ser  $\geq 2,4$ ” presenta determinados aspectos estadísticos relevantes, especialmente en entornos móviles reales:

#### 3.1 Sensibilidad a tamaños muestrales

Con el valor muestral mínimo planteado en las fichas metodológicas del Proyecto Normativo, una pequeña variación puede cambiar artificialmente el resultado entre cumplimiento e incumplimiento.

#### 3.2 Alta influencia de valores atípicos (outliers)

Eventos específicos como: Handovers sucesivos, CSFB (Circuit Switched Fallback), SRVCC (Single Radio Voice Call Continuity), interferencias puntuales, condiciones de movilidad extrema pueden introducir muestras degradadas no representativas de la experiencia general.

#### 3.3 Naturaleza no gaussiana del MOS

Las distribuciones de MOS en drive test suelen ser asimétricas y multimodales. En estas condiciones, los conteos simples o promedios pierden robustez.

### 4. Fundamentación estadística del uso del Percentil

La Recomendación UIT-T E.840, en la Sección 9.3.1 “Unidad de calidad de funcionamiento estadística” (página 12), establece que los análisis comparativos pueden basarse en unidades estadísticas que reflejen la calidad media de la red o su coherencia, representada por la P-ésima probabilidad de estar por encima de un valor umbral predefinido (UIT-T E.840, 06/2018):

#### **9.3.1 Unidad de calidad de funcionamiento estadística**

Los análisis comparativos deben basarse en unidades de calidad de funcionamiento estadísticas que reflejen la calidad de funcionamiento media de la red (representada por valores medios,  $m$ ) o su coherencia (representada por la P-ésima probabilidad de estar por encima de un valor umbral predefinido). En esta Recomendación se utiliza como ejemplo la calidad de funcionamiento estadística media. Otras técnicas semejantes pueden aplicarse para la coherencia.

En este sentido, la propia recomendación de la UIT, citada por la misma ARCOTEL en su Informe dice explícitamente que una unidad estadística válida para analizar calidad puede ser:

- La media
- La P-ésima probabilidad/percentil de estar por encima de un umbral

Esto valida formalmente el uso de percentiles como herramienta adecuada para evaluar cumplimiento de MOS.

#### 4.1 Ventajas técnicas del Percentil 90

El Percentil 90 (P90) presenta ventajas claras frente a métricas de cumplimiento porcentual:

- No requiere supuestos de normalidad
- Reduce la influencia de valores extremos

- Representa la experiencia de la gran mayoría de usuarios
- Es estable ante variaciones puntuales de red

## 5. Consistencia con las recomendaciones internacionales (ITU y ETSI)

Las normas y reportes técnicos internacionales utilizados por ARCOTEL no solo permiten, sino que recomiendan explícitamente el uso de percentiles:

### ETSI TS 102 250-6

Establece que, para parámetros de calidad perceptual, los cuantiles (percentiles) son más interpretables que los promedios:

**8 Data aggregation**  
**8.1 Basic data aggregation operators**  
*Depending on the objective, i.e. for monitoring or reporting purposes, different types of aggregation may be of interest. Firstly, different granularities in time and space may be needed. Secondly, weighting of data may be considered for instance to generate the QoS perceived by the user. This may be more or less straightforward, if the smallest time granularity is evenly spaced and the full information at any time point or interval is available, i.e. there are no missing data. However, data aggregation becomes more challenging if event data are considered, like data from active probing systems, or if data are missing and substitution algorithms are needed for defining meaningful results at higher aggregation levels.*  
*In the following, after presenting some basic aggregation operators, different data sources and corresponding structures are distinguished, temporal and spatial aggregation levels are defined and estimation methods for incomplete data are suggested. Based on a summary of desirable attributes of aggregation procedures, an aggregation strategy is suggested and discussed. Subsequently, weighted aggregations are motivated and weighting methods are introduced.*  
*Most common data aggregation operators are sums and means. Sums are applied if the total number of events, usage, etc. within a given time period is of interest, while means are commonly applied if some kind of averaging is needed. However, in particular for data out of time-interval measurements, means may not lead to interpretable and convincing aggregation results, therefore other summarizing statistics like the minimum, maximum, median and other quantiles of data are also used as aggregation operators. For quality of*

### ETSI TR 103 559 (Benchmarking QoS/QoE)

Utiliza explícitamente Percentil 90 de MOS y Ratios de MOS por debajo de umbrales, como métricas estándar de evaluación comparativa, no usa un porcentaje de muestras de cumplimiento, tal como afirma erróneamente el Informe de Presentación del Proyecto Normativo de ARCOTEL.

**9.3.2 Telephony**  
 (...)  
**9.3.2.2 Scoring**  
*The higher the call success ratio and the MOS value the better is the experience of the user; the two values have an increasing value score. While the longer the call setup time is, the worse the experience of the user is; the call setup time has a decreasing score value. In addition to the consideration of the average values, extra bonus for excellent listening quality or very short call setup-times can be given, the same as extra bonus for the reduction of very bad experiences as very low MOS scores or very long call setup times. As examples, the 10th percentile of CST can be used for awarding very short CST and the 90th percentile for awarding excellent listening quality. To award the absence of negative experiences the 90th percentile of CST can be considered and a ratio of MOS < threshold (e.g. MOS < 1,6 or MOS < 2,2). A lower parameter value would lead to a higher score in these cases.*  
*For telephony the following factors with example thresholds and percentiles can be taken into account:*

- Call Setup Success Ratio.
- Call Drop Ratio.
- MOS.
- MOS < low MOS threshold.
- 90th percentile of MOS.
- Call Setup Time.
- Call Setup Time > long setup time threshold.
- 90th percentile of Call Setup Time.

## UIT-T E.840

Reconoce la agregación mediante percentiles como método válido de evaluación de calidad.

### **9.3.1 Unidad de calidad de funcionamiento estadística**

Los análisis comparativos deben basarse en unidades de calidad de funcionamiento estadísticas que reflejen la calidad de funcionamiento media de la red (representada por valores medios,  $m$ ) o su coherencia (representada por la P-ésima probabilidad de estar por encima de un valor umbral predefinido). En esta Recomendación se utiliza como ejemplo la calidad de funcionamiento estadística media. Otras técnicas semejantes pueden aplicarse para la coherencia.

En consecuencia, el uso del “Percentil 90” no es una flexibilización, sino una aplicación fiel de las mejores prácticas internacionales.

## **6. Interpretación desde la experiencia humana**

Desde el punto de vista perceptual:

- El cerebro humano tolera breves degradaciones si la mayor parte de la conversación es nítida.
- Si el 90 % del tiempo la conversación es inteligible, la experiencia global se percibe como aceptable.
- Pretender que el 100 % de las muestras superen el umbral no es técnicamente viable en entornos radioeléctricos móviles.

El “Percentil 90” introduce una tolerancia estadísticamente justa (10 %), alineada con:

- La física del canal radio,
- La movilidad del usuario,
- La variabilidad propia del entorno urbano y rural.

## **7. Propuesta técnica de CONECEL**

En función del análisis expuesto, mi Representada propone que el parámetro Calidad de Conversación (MOS) sea evaluado mediante el “Percentil 90” del MOS ( $P_{90}$ )  $\geq 2,4$ ; como metodología técnica y estadísticamente robusta y alineada con los estándares internacionales adoptados por la propia ARCOTEL en el Informe de Presentación del Proyecto Normativo.