# PEMO-Q

AudioQual, AudioQual_HI and SpeechQual

Audio and speech quality prediction for normal hearing and hearing impaired people, based on the Oldenburg Perception Model (PEMO)

Version 1.4.1

*Manual*

# Table of Contents



**Note:**

# License Agreement

**IMPORTANT NOTE:**

**Read the following license agreement carefully before you install or use the Software.**

**By installing the Software (as defined below), copying the Software and/or using the Software, you (either on behalf of yourself as an individual or on behalf of an entity as its authorized representative) agree to all of the terms of this end user license agreement ("agreement") regarding your use of the Software.**

**If you do not agree with all of the terms of this agreement, do not install, copy or otherwise use the Software.**

## End User License Agreement

DEFINITIONS

The term "Software" includes all software distributed with this License including all documentation. The "Software" is licensed to you under the terms specified in the License Grant below.

HIGH RISK ACTIVITIES

The Software is not fault-tolerant and is not designed, manufactured or intended for use as on-line control equipment in hazardous environments requiring fail-safe performance, such as in the operation of nuclear facilities, aircraft navigation or communication systems, air traffic control, direct life support machines or other medical devices, or weapons systems, in which the failure of the Software could lead directly to death, personal injury, or severe physical or environmental damage ("High Risk Activities").

The University of Oldenburg and its suppliers specifically disclaim any express or implied warranty of fitness for High Risk Activities.

OWNERSHIP AND COPYRIGHT

This Software is owned by the University of Oldenburg and is protected by copyright law and international copyright treaty. Therefore you must treat this Software like any other copyrighted material. You acknowledge that no title to the intellectual property in the Software is transferred to you. Title, ownership, rights, and intellectual property rights in and to the Software shall remain in the University of Oldenburg.

LICENSE GRANT

Subject to the license terms, the University of Oldenburg hereby grants you a non-exclusive, non-transferable license to install and to use the Software under the terms of this license. Except as provided in this license agreement, **you may not sell, rent, lease, modify, translate, or sublicense the Software**. You agree not to modify the Software or attempt to decipher, de-compile, disassemble or reverse engineer the Software, except to the extent applicable laws specifically prohibit such restriction.

## DISCLAIMER OF WARRANTY

This Software is provided and/or sold "as is" and without warranties as to performance of merchantability or any other warranties whether expressed, implied, or statutory, including, but without limitation, the implied warranties of noninfringement of third party rights, merchantability, and fitness for a particular purpose. Good data processing procedure dictates that any program be thoroughly tested with non-critical data before relying on it. The user must assume the entire risk of using the program. Any liability of the seller will be limited exclusively to product replacement or refund of purchase price.

Under and restricted by the above terms, the University of Oldenburg warrants that the software, as updated and when properly used, will perform substantially in accordance with its accompanying documentation, and the Software media will be free from defects in materials and workmanship.

The limited warranty is void if the software fails as a result of accident, abuse, misapplication or modification.

## LIMITATION OF LIABILITY

You must assume the entire risk of using the Software.

**In no event shall the University of Oldenburg be liable to you for any damages, including any lost profits, lost savings, or other incidental, indirect or consequential damages of any kind arising out of the use of the University of Oldenburg's Software, even if the University of Oldenburg has been advised of the possibility of such damages. In no event will the University of Oldenburg's liability for any claim, whether in contract, tort, or any other theory of liability, exceed the license fee paid by you. This limitation shall apply to claims of personal injury to the extent permitted by law.**

## System Requirements

The software „PEMO-Q AudioQual, AudioQual_HI and SpeechQual" (audio and speech quality prediction based on the Oldenburg Perception Model) version 1.4.1 ("the software") is provided as **MATLAB® functions** (as pre-parsed pseudo-code files [P-files] with accompanying mex files) under MS Windows® operating systems, LINUX and macOS. The respective function names are "audioqual", "audioqual_hi" and "speechqual", respectively.

## Installation

To install the software on your computer, simply copy all files into a directory of your choice on your computer.

Start MATLAB® and change to the installation directory of the software or add that directory to your MATLAB® path. You can now call the functions `audioqual`, `audioqual_hi` and `speechqual` according to the syntax described in the functions' help texts (→ type "`help audioqual_help`", "`help audioqual_hi_help`" or "`help speechqual_help`", respectively). (See also section "Usage and description" of this manual for information on the usage of these functions.)

## About PEMO-Q; recommendations for usage

The present PEMO-Q software package serves to predict perceived quality differences between audio signals. It is based on the work of Huber and Kollmeier (2006), which in turn follows the approach by Hansen and Kollmeier (2000), using the "Oldenburg Perception Model" ("PEMO") of Dau et al. (1996, 1997) for computing "internal representations" of signal pairs. These pairs are compared quantitatively basically by calculating the linear cross correlation coefficient. The resulting correlation value serves as an objective measure of the perceptual similarity between two audio signals.

The method of Hansen and Kollmeier was originally developed to predict the quality of telephone-band limited speech with relatively distinct distortions caused by low-bit-rate speech codecs. This approach is reproduced in the module SpeechQual of the software package.

The module AudioQual is used to predict quality degradations of general broad-band audio signals. The sensitivity for minor quality distortions can be increased by selecting a filter bank instead of a low-pass filter for modulation processing in the perception model (see "Usage and description"). However, this also increases the computation time considerably. Apart from the overall correlation between internal representations (output value 'PSM'), AudioQual also computes an estimate of the instantaneous audio quality as a function of time by frame-wise correlation (output vector 'PSM_inst'). The non-linear average of this times series represents another estimate of the perceived overall quality (output value 'PSMt' / 'ODG'). The latter, more complex measure has especially been designed and optimized to achieve good prediction results for the evaluation of audio codecs (Huber and Kollmeier, 2006). On the other hand, recent studies, where PEMO-Q has been applied in fields other than audio codec evaluation, e.g. speech enhancement algorithm evaluation, indicate a more general applicability of the "simpler" measure PSM. Except for the assessment of very small quality distortions, e.g., in the evaluation of high-quality audio codecs, a general better performance of the modulation filter-bank version compared to the much faster low-pass version has not been found. Especially in case of rather distinct quality differences, the higher sensitivity of the filter-bank version seems not to yield any significant advantage which would justify the higher computational effort.

# Usage and description

## AudioQual

### USAGE[1]

MATLAB® call syntax:

```
[PSM, PSMt, ODG, PSM_inst] = audioqual(RefSig, TestSig, fs, dt, modproc, MaxLag,
PauseCut, LevelAlign, assim, verbose)
```

### Input arguments:

```
   RefSigFile : reference signal (either vector or wav-filename) (*)
  TestSigFile : test signal (either vector or wav-filename) (*)
           fs : sample rate (optional if signals are specified by wav-filenames)
           dt : time resolution (correlation interval length) in ms
                (optional, default = 10)
      modproc : type of modulation processing in PEMO:
                either filterbank ("fb") or lowpass ("lp") (default: "lp")
                "fbN" with 1 <= N <= 10 applies a modulation filterbank with N
                filters (optional, default: N = 8)
       MaxLag : parameter for automatic delay compensation. MaxLag defines the range
                of lags in ms ([-MaxLag, MaxLag]), over which the search for a
                possible delay between reference and test signal is performed.
                Set = 0 to skip the delay compensation(**). (optional, default = 0)
     PauseCut : flag for removing silent intervals (i.e. -70 dB re. full scale in
                RefSigFile) of at least 200 ms duration from signals
                (optional, default = 0)
   LevelAlign : flag for compensating a possible overall level difference
                (optional, default = 1)
        assim : flag for partial assimilation of internal representations
                (Beerends-Berger-approach) (optional, default = 1)
      verbose : flag for displaying parameter settings and progress messages
                (optional, default = 0)
```

(Keep order of arguments, e.g. if you want to specify *modproc*, also specify *dt*; you may type "[]" to adopt default values.)

(*) A (square wave) signal amplitude of 1 is assumed to correspond to 100 dB SPL.

(**) If the delay is known, it is recommended to align the signals beforehand and skip the automatic delay compensation.

### Output arguments:

```
   PSM : "Perceptual Similarity Measure": 1st overall objective quality
         measure (= overall correlation between internal representations)
  PSMt : 2nd overall objective quality measure
```

[1] Type "`help audioqual_help`" to get this help text.

```
            (= 5th percentile of "internal activity"-weighted PSM_inst)
       ODG : "Objective Difference Grade", i.e transformed PSMt measure
  PSM_inst : vector of instantaneous objective quality
            (i.e. PSM_inst = PSM(t) with t = n*dt ms, n = 1,2,...; def.: dt = 10)
```

**DESCRIPTION**

AudioQual is an implementation of the method for objective perceptual assessment of audio quality "PEMO-Q" (Huber and Kollmeier, 2006; slightly modified). It predicts the perceived audio quality of a given test signal relative to that of a reference signal, using the model of auditory perception ("PEMO") by Dau et al. (1996, 1997). After pre-alignment of the given pair of audio signals (delay compensation and level alignment), the auditory model transforms both signals into corresponding internal representations. Following an approach of Beerends/Berger, an optional subsequent assimilation procedure reduces the differences between the internal representations sign-dependently, putting more weight on positive than on negative deviations of the processed signal. Finally, the linear cross correlation coefficient of the pair of internal representations, PSM, is calculated and serves as a measure of the perceived similarity. If the reference signal is of high audio quality, PSM may also be interpreted as an objective measure of the audio quality (degradation) of the test signal. In addition, a time series of instantaneous audio quality is computed by consecutively correlating slices of the internal representations (length of the slices = *dt* ms, 10 ms per default). The output vector PSM_inst contains the sequence of PSM(n**dt*), n = 1, 2, …. The overall quality measure PSMt is calculated by taking the 5th percentile of PSM_inst, after it is weighted by the moving average of the internal representation of the test signal ("internal activity" weighting). The third output argument, ODG, is obtained by mapping PSMt to a value that corresponds to the "Subjective Difference Grade" (SDG) quality scale defined in ITU-R BS.1116. The mapping function *f*: PSMt → ODG was derived empirically by fitting the results of several listening tests performed for the evaluation of audio codecs. THIS MAPPING ONLY HOLDS IF DEFAULT SETTINGS OF *dt, LevelAlign, assim* AND WIDE BAND SIGNALS ARE USED! (Regarding *modproc*, settings "lp", "fb" and "fb8" are possible.) Moreover, be aware that this mapping might also not hold for applications other than audio codec evaluation.

The values of SDG and ODG have the following meanings:

The quality impairment of the test signal relative to the reference signal is

 0 : Imperceptible
-1 : Perceptible but not annoying
-2 : Slightly annoying
-3 : Annoying
-4 : Very annoying

Reference:

Huber, R. and Kollmeier, B. (2006). "PEMO-Q–A new Method for Objective Audio Quality Assessment using a Model of Auditory Perception." IEEE Transactions on Audio, Speech and Language processing, Vol. 14, no. 6, pp. 1902 - 1911.

## AudioQual_HI

### USAGE[2]

MATLAB® call syntax:

[*PSM, PSMt, ODG, PSM_inst*] = audioqual_hi(*RefSig, TestSig, fs, HL, dt, modproc, MaxLag, PauseCut, LevelAlign, assim, verbose*)

### Input arguments:

```
  RefSigFile : reference signal (either vector or wav-filename) (*)
 TestSigFile : test signal (either vector or wav-filename) (*)
          fs : sample rate (optional if signals are specified by wav-filenames)
          HL : Hearing loss, i.e. audiogram data: 2xN matrix;
               first row: frequencies in Hz, second row: pure tone hearing
               threshold in dB HL
               set HL=0 or HL=[] to indicate normal hearing (optional, default: 0)
          dt : time resolution (correlation interval length) in ms
               (optional, default = 10)
     Modproc : type of modulation processing in PEMO:
               either filterbank ("fb") or lowpass ("lp") (default: "lp")
               "fbN" with 1 <= N <= 10 applies a modulation filterbank with N
               filters (optional, default: N = 8)
      MaxLag : parameter for automatic delay compensation. MaxLag defines the range
               of lags in ms ([-MaxLag, MaxLag]), over which the search for a
               possible delay between reference and test signal is performed.
               Set = 0 to skip the delay compensation(**). (optional, default = 0)
    PauseCut : flag for removing silent intervals (i.e. -70 dB re. full scale in
               RefSigFile) of at least 200 ms duration from signals
               (optional, default = 0)
  LevelAlign : flag for compensating a possible overall level difference
               (optional, default = 1)
       assim : flag for partial assimilation of internal representations
               (Beerends-Berger-approach) (optional, default = 1)
     verbose : flag for displaying parameter settings and progress messages
               (optional, default = 0)
```

(Keep order of arguments, e.g. if you want to specify *modproc*, also specify *dt*; you may type "[]" to adopt default values.)

(*) A (square wave) signal amplitude of 1 is assumed to correspond to 100 dB SPL.

(**) If the delay is known, it is recommended to align the signals beforehand and skip the automatic delay compensation.

### Output arguments:

```
      PSM : "Perceptual Similarity Measure": 1st overall objective quality
            measure (= overall correlation between internal representations)
     PSMt : 2nd overall objective quality measure
            (= 5th percentile of "internal activity"-weighted PSM_inst)
      ODG : "Objective Difference Grade", i.e transformed PSMt measure
 PSM_inst : vector of instantaneous objective quality
            (i.e. PSM_inst = PSM(t) with t = n*dt ms, n = 1,2,...; def.: dt = 10)
```

---

[2] Type "help audioqual_hi_help" to get this help text.

### DESCRIPTION

audioqual_hi corresponds to audioqual, extended by an instantaneous expansion stage after the haircell stage in the auditory model PEMO, as proposed by Derleth et al. (2001), to model effects of sensorineural hearing loss.

Reference:

Huber, R., Parsa, V., and Scollie, S. (2014) "Predicting the perceived sound quality of frequency-compressed speech," PLoS ONE 9(11):e110260.
DOI: 10.1371/journal.pone.0110260

## SpeechQual

### USAGE[3]

MATLAB® call syntax:

```
qc = speechqual(RefSig, TestSig, fs, MaxLag, PauseCut, LevelAlign, weighting,
assim, verbose)
```

**Input arguments**:

```
     RefSig :  reference signal (either vector or wav-filename) (*)
    TestSig :  processed (test) signal (either vector or wav-filename) (*)
         fs :  sample rate
     MaxLag :  parameter for automatic delay compensation. MaxLag defines the range
               of lags in ms ([-MaxLag, MaxLag]), over which the search for a
               possible delay between reference and test signal is performed. Set =
               0 to skip the delay compensation(**). (optional, default = 0)
   PauseCut :  flag for removing silent intervals of at least 200 ms from signals
               (optional, default = 0)
 LevelAlign :  flag for compensating a possible overall level difference (optional,
               default = 1)
  weighting :  flag for Hansen's band importance weighting (optional, default = 1)
      assim :  flag for partial assimilation of internal representations (Beerends-
               Berger-approach) (optional, default = 0)
    verbose :  flag for displaying parameter settings and progress messages
               (optional, default = 0)
```

**Output arguments:**

```
  qc : objective speech quality measure
```

### DESCRIPTION

SPEECHQUAL is an implementation of Hansen's method for speech quality estimation. It predicts the perceived quality of a given test speech signal relative to that of a reference signal, using the model of auditory perception ("PEMO") by Dau et al. (1996). After prealignment of the given pair of speech signals, the auditory model transforms both signals into corresponding internal representations. A set of linear weights is applied to the frequency bands of the two representations to account for different assumed importances for the perceived speech quality. (Can be disabled.) Following

[3] Type "`help speechqual_help`" to get this help text.

an approach of Beerends/Berger, an optional subsequent assimilation procedure (not contained in Hansen's original method) can be applied to reduce the differences between the internal representations sign-dependently, putting more weight on positive than on negative deviations of the processed signal (default: no assimilation). Finally, the linear cross correlation coefficient of the pair of internal representations, "qc", is calculated and serves as a measure of the perceptual similarity. If the reference signal is of high quality, qc may be interpreted as an objective measure of the speech quality (degradation) of the test signal.

Reference:

Hansen, M. and Kollmeier, B. (2000). "Objective modelling of speech quality with a psychacoustically validated auditory model," J. Audio Eng. Soc., vol. 48(5), 395-409

# References


Huber, R. and Kollmeier, B. (2006). "PEMO-Q–A new Method for Objective Audio Quality Assessment using a Model of Auditory Perception." IEEE Transactions on Audio, Speech and Language processing, Vol. 14, no. 6, pp. 1902 - 1911.

Huber, R. (2003). "Objective assessment of audio quality using an auditory processing model," Dissertation, Universität Oldenburg, BIS, http://oops.uni-oldenburg.de/182/1/hubobj03.pdf

Huber, R., Parsa, V., and Scollie, S. (2014) "Predicting the perceived sound quality of frequency-compressed speech," PLoS ONE 9(11):e110260. DOI: 10.1371/journal.pone.0110260

Hansen, M. and Kollmeier, B. (2000). "Objective modelling of speech quality with a psychacoustically validated auditory model," J. Audio Eng. Soc., vol. 48(5), 395-409

Hansen, M. (1998). "Assessment and prediction of speech transmission quality with an auditory processing model," Dissertation, Universität Oldenburg, BIS

Dau, T., Püschel, D., and Kohlrausch, A. (1996). "A quantitative model of the 'effective' signal processing in the auditory system. I. Model structure," J. Acoust. Soc. Am 99(6), 3615-3622

Dau, T., Kollmeier, B., and Kohlrausch, A. (1997). "Modeling auditory processing of amplitude modulation." J. Acoust. Soc. Am. 1997, vol. 102, no. 5, p. 2892 – 2905

Derleth, R.P., Dau, T., and Kollmeier, B. (2001). Modelling temporal and compressive properties of the normal and impaired auditory system. Hearing Research 159, 132-149.