

DGA 2014 tutorial 6

“Auditory models and their application”

Tim Jürgens, Ray Meddis, Carolin Iben

Saturday, March 15, 2014

Content

[1. Model of the auditory periphery (MAP) 3](#_Toc379114057)

[1.1 Introduction 3](#_Toc379114058)

[1.2 Block diagram 3](#_Toc379114059)

[1.3 Exercises 3](#_Toc379114060)

[1.3.1 Input signal 3](#_Toc379114061)

[2. Computational auditory signal processing and perception (CASP) model 4](#_Toc379114062)

[2.1 Introduction 4](#_Toc379114063)

[2.2 Block diagram 4](#_Toc379114064)

[2.3 Exercises 6](#_Toc379114065)

[2.3.1 Getting started **Fehler! Textmarke nicht definiert.**](#_Toc379114066)

[2.3.2 From input signal to internal representation 6](#_Toc379114067)

[2.3.2 Predict and evaluate thresholds **Fehler! Textmarke nicht definiert.**](#_Toc379114068)

[3. Binaural speech intelligibility model (BSIM) 7](#_Toc379114069)

[3.1 Introduction 7](#_Toc379114070)

[3.2 Block diagram 7](#_Toc379114071)

[3.3 Exercises 7](#_Toc379114072)

[3.3.1 Input signal 7](#_Toc379114073)

# Model of the auditory periphery (MAP)

## Introduction

MAP1\_14.m is the latest version of the Matlab Auditory Periphery (MAP) model from Essex University, UK. It is a computer program that simulates all stages of the auditory periphery from the tympanic membrane up to the auditory nerve (AN) and into the brainstem. A special feature of the model is the close modelling of the physiology of the periphery.

MAP1\_14 now has efferent functionality including both acoustic reflex and MOC efferent function. These are recent additions and remain experimental in terms of their details.

The software can be downloaded from

<https://dl.dropboxusercontent.com/u/13144068/MAP1_14h.zip>

The software package contains extensive documentation in the ‘documentation’ folder including a WORD document called ‘MAP1\_14 Quick Start’.

The model is realised by systems of equations that are described in a second document called ‘MAP Model Technical Description’.

MAP1\_14 can be used to:

1. Process acoustic waveforms (e.g. tones, speech) to generate multi-channel representations at different levels in the auditory periphery including tympanic membrane, stapes, basilar membrane (BM) displacement, inner hair cell cilia displacement, receptor potential and auditory nerve. Additional representations are possible at brainstem level in terms of neuronal responses including chopper and primary-like responses in the cochlear nucleus (CN). Second level neurons can also be simulated.
2. Illustrate basic physiology experiments in the auditory periphery.
3. Generate auditory features for automatic speech recognisers.
4. Demonstrate a range of psychophysical phenomena such as absolute threshold, threshold as a function of duration, forward masking, threshold masking curves (TMCs), psychophysical tuning curves (PTCs).

## Installation

The software comes in a zipped folder called MAP1\_14.

Unzip the folder and place it anywhere in the main memory. No further installation is required

## Block diagram

The model is implemented as a series of processing stages in the middle and inner ear. The response at each stage is computed and can be separately displayed. Two feedback loops representing the MOC and acoustic reflexes have been added. These are the focus of current research efforts.

|  |
| --- |
|  |

Figure 1: Schematic of the auditory pathway modelled by MAP1\_14. Each shape represents a separate stage of signal processing in the auditory periphery.

## Exercises

### demoClick

After launching MATLAB, set the current folder to MAP1\_14h\demonstrations.

In the command window type ‘run demoClick’. The program should run and complete in a few seconds. This model simulates the response to a click stimulus of the stapes and 21 locations along the basilar membrane (BM) with BFs between 250 and 4000 Hz before computing the firing probability for high spontaneous rate (HSR) auditory nerve (AN) fibers in the corresponding channels.

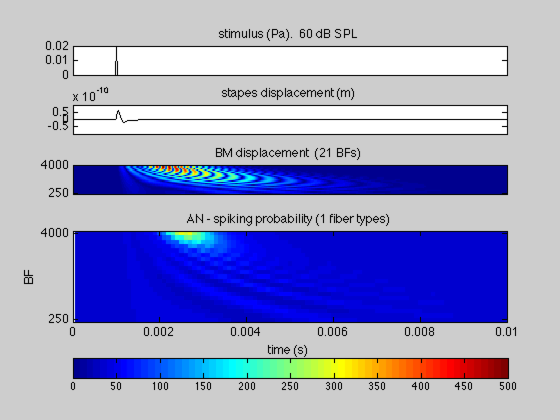
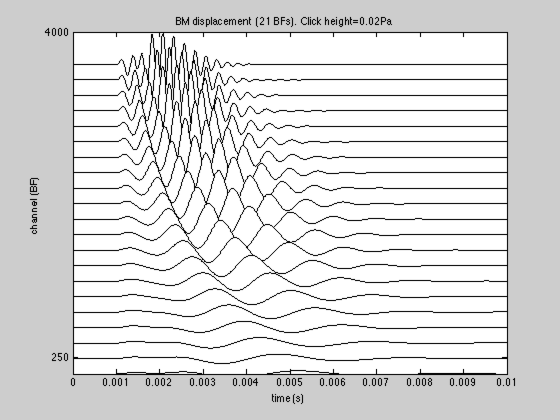
 

Figure 2. demoClick. A: processing stages (stimulus, stapes displacement, BM displacement and AN firing probability). B: cascade plot showing the vibration of the BM at each location.

### demoTwisterProbability

A second demonstration can also be found in the ‘demonstrations’ folder. In the command window type

demoTwisterProbability

This demonstration uses a wav file recording of the word ‘Twister’ as the input to the program (replacing the click in the example above) and shows the response at different auditory processing stages. The AN(HSR) firing probability is shown as a surface plot in the second image.

|  |  |
| --- | --- |
|  |  |

Figure 3: **Left panel: Model output for the demonstration file *demoTwisterProbability*. The input utterance is the word ‘twister’ spoken at an rms level of 60 dB SPL by a male speaker. The model has 21 channels with BFs between 250 and 8000 Hz (evenly spaced on a logarithmic scale). The top three panels represent the signal, the stapes displacement and the basilar membrane displacements respectively. The bottom panel shows the AN response firing probability computed for high spontaneous rate fibers. Right panel: Surface plot of the high-spontaneous rate fiber response. This plot represents the probability of action potentials scaled up to represent firing rates in each BF channel.**

### demoTwisterSpikes

This demonstration is similar to demoTwisterProbability except the AN response is represented as individual stochastic spikes rather than spike probability. For this reason, each spike is shown as a separate dot in the figure (4th panel down). When the AN spiking activity is modelled as individual spikes, it is possible to continue processing into the brainstem. The bottom two panels simulate spiking activity in two representative sets of neurons at the levels of

1. cochlear nucleus (CN, sustained chopping neurons). These receive inputs from the AN.

2. ventral nucleus of the trapezoid body (VNTB). These neurons receive inputs from the CN. Their activity is used to control the MOC suppressive feedback loop to the BM.

MOC efferent feedback is used only in advanced applications and will not be discussed further. It is included here for those who are curious about this feature.

|  |
| --- |
|  |

Figure 4: **Model response to the utterance ‘twister’. The first three panels are the same as the previous figure (signal, stapes response, BM response). The auditory nerve response (panel 4) is given as a raster plot of spike activity in 21 channels for high spontaneous rate fibers (2100 fibers altogether). The bottom two panels show the spiking activity of simulated chopper units in two successive layers of brainstem activity. All units are programmed as simple chopper cells.**

### Further exercises

More examples and exercises can be found in the document ‚ ‘MAP1\_14 Quick Start’ in the ‚’documentation’ folder

# Perception model

## Introduction

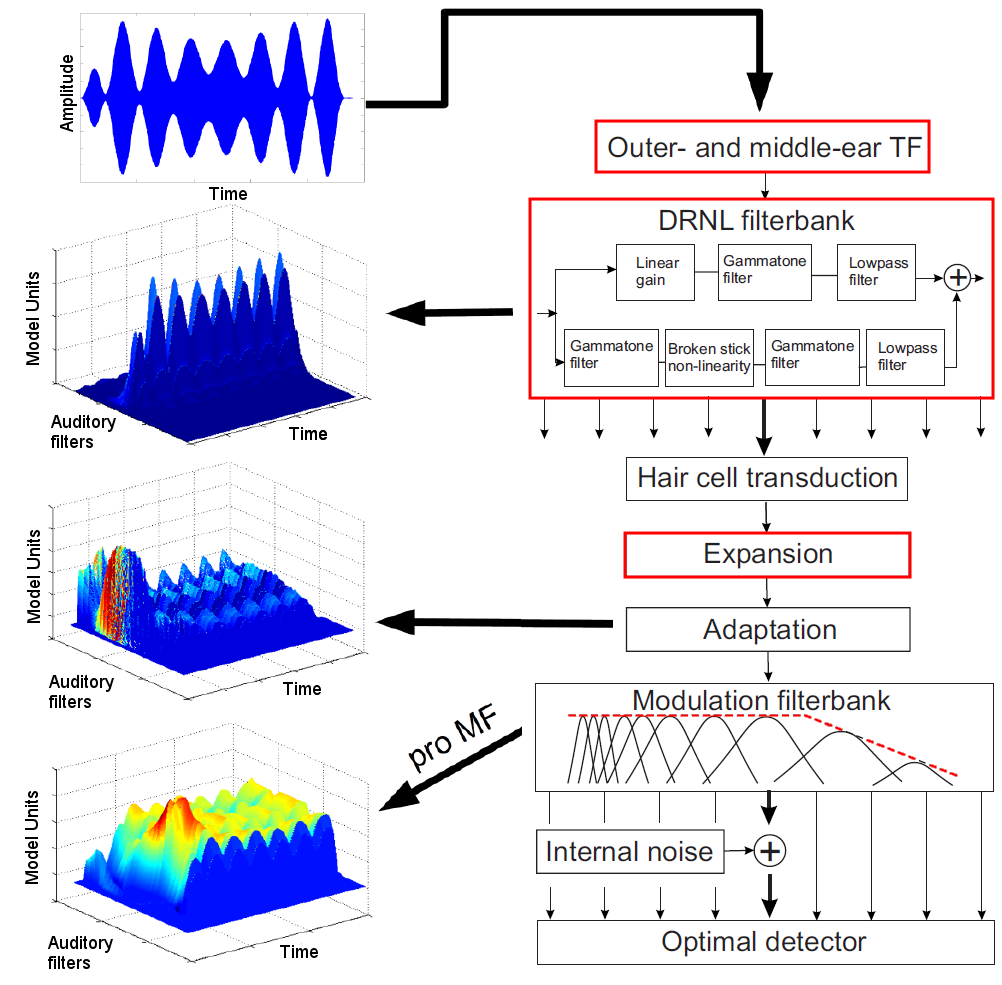
The “perception model” is a model of the effective auditory signal processing of normal hearing listeners. The term “effective” refers to the fact that while the processing steps are inspired from knowledge about the underlying physiology, they are implemented as simplified signal processing steps covering the essentials. Two models can be uses in the current exercise, the “original” perception model PEMO by Dau et al. (1997a) and the successor computational model of human auditory signal processing and perception (CASP) by Jepsen et al. (2008). Both models were shown to account for a large variety of psychoacoustic masking and signal detection experiments.

The goal of this exercise is to gain insight in the auditory preprocessing stages of the models which transform the physical stimulus waveform to the so called „internal representation“. The internal representation of the model is meant to contain all signal features that are relevant for auditory perception. When psychoacoustic experiments are simulated with the model, the internal representation of a stimulus, e.g. a tone in noise, is compared with the „learned“ internal representation of the tone alone. This is done in the detector stage of the model which follows the proprocessing. The detector stage is not content of this exercise.

## Block diagram

The block diagram shows the stages of the CASP model on the right and their exemplary outputs on the left. The red highlighted components are different in the CASP compared to the PEMO model. The stages can be separated in the auditory preprocessing and the detector stage. The preprocessing of CASP consists of the following stages with their respective functions:

* Outer and middle-ear transfer function: Frequency shaping according to the transduction through ear canal and middle-ear bones. This stage is not contained in PEMO.
* Basilar-membrane filtering (DRNL; dual resonance non-linear filter): Decomposition of the input in different frequency bands (typically around 30) including level dependent filter and compression characteristics. In PEMO a linear Gammatone filterbank is used.
* Haircell transduction and expansion: Half-wave rectification and lowpass filtering (1 kHz) to mimic the auditory systems ability to follow fine structure at low frequencies and the envelope at higher frequencies. In CASP a squaring expansion follows.
* Adaptation: The first key element of PEMO/CASP mimicking emphasis of signal onsets. Moreover mean of the input (equivalent to the mean level) is logarithmically compressed while faster fluctuations (amplitude modulations above about 4 Hz) are linearly processed. The logarithmic compression of the mean level enables the prediction of Weber’s law for intensity discrimination.
* Modulation filtering: The second key element of PEMO/CASP accounting for frequency selective processing of amplitude modulations.



## Exercises

Unzip ASPMT\_DGA2014.zip to your Matlab directory. Change to the folder \ASPMT\_DGA2014.

### 2.3.1 Input signal

The input signal can be any wav file. Some examples are provided in the \demo folder. The default input signal (x.wav) is a 3-Hz wide 5-kHz pure tone amplitude modulated at 20 Hz, similar to the input used in the block diagram. To see the time waveform type in the command window

DemoMain(‘casp’,’input’).

Due to the narrow bandwidth a slow variation in the signal amplitude is visible in addition to the 20-Hz fluctuation. The fine structure is not resolved in the plot without zooming in.

### 2.3.2 From input signal to internal representation

The output of all stages is shown by DemoMain(‘casp’,’all’) or step by step by stating DemoMain(‘casp’,’pmode’) with pmode = ‘input’,’basilar’,’hc’,’adapt’. Instead of ‘casp’ ‘pemo’ can be used. For more details browse: DemoMain.m or key in help DemoMain in the command window.

Because the final internal representation is a three dimensional time-varying activity pattern, only a two-dimensional activity pattern across auditory and modulation filters is plotted for the last stage with energy in color coding. Here the concentration of activity in the 5-kHz region and at 20-Hz modulation rate is visible. The peak at the lowest modulation filter (a lowpass filter) resembles the signal level. To view this stage type DemoMain(‘casp’,’intRepImage’).

The internal representation per auditory filter, corresponding to the time signal at the output of the modulation filters per auditory filter is produced by DemoMain(‘casp’,’intRep’). For more details browse scripts\PlotIntRepMain or type help PlotIntRepMain in the command window.

* Model parameters like the frequency ranges of the filterbanks can be adjustted in demo\Demo\_casp\_cfg.m and demo\Demo\_pemo\_cfg.m which can be found in the \demo folder.
* Different input signals are provided to assess the effect of signal transformation stages: TODO

### 2.3.3 Further reading:

.....dsaf

# Binaural speech intelligibility model (BSIM)

## Introduction

## Block diagram

## Exercises

### Input signal