An Algorithm for Statistical Audibility Prediction (SAP) of an Arbitrary Signal in the Presence of Noise

Menachem Rafaelof
National Institute of Aerospace at NASA Langley Research Center

Contributors
Andrew Christian and Matthew Boucher
NASA Langley Research Center

Audio Engineering Society (AES)
153rd Convention, New York
October 19-20, 2022
Workshop #13

Outline

- The need for predicting audibility
- Challenges and Basic Approach
- SAP
  - Basics (specific loudness and d-prime)
  - Algorithm steps
- Performance validation at low frequencies
- Ongoing
  - Hearing below the level of the masker, high frequency signals and binaural hearing
- Demonstration and a task for this audience
  - Helicopter over NYC
  - SUI drone over UPS truck
  - Distributed Electric Propulsion (DEP) vehicle over box truck

M. Rafaelof and K. Wendling, "An Algorithm for Statistical Audibility Prediction (SAP) of an Arbitrary Signal in the Presence of Noise" *J. Audio Eng. Soc.*, vol. 69, no. 9, pp. 672–682, (2021 September). DOI: https://doi.org/10.17743/jaes.2021.0021

NASA LAR-19736-1: patent pending

The need for predicting audibility

Avoid intrusive or annoying sounds

Unmanned Arial vehicles (UAs)

Urban Air Mobility (UAM) vehicles

Other sources

Assure audibility

Alarms, telephone ring tone, speech, music over a given ambient noise. etc.

Some benefits:

- A common method for virtual assessment of any sound in presence of another sound
- Explaining root cause for audibility vs. time and/or frequency
- **As an effective noise metric (hot topic – e.g., AES standards SC-04-09)**

Challenges and Basic Approach

A flexible algorithm to predict the audibility of a given signal in presence of masker.

Deal with any combination of a signal and a masker:

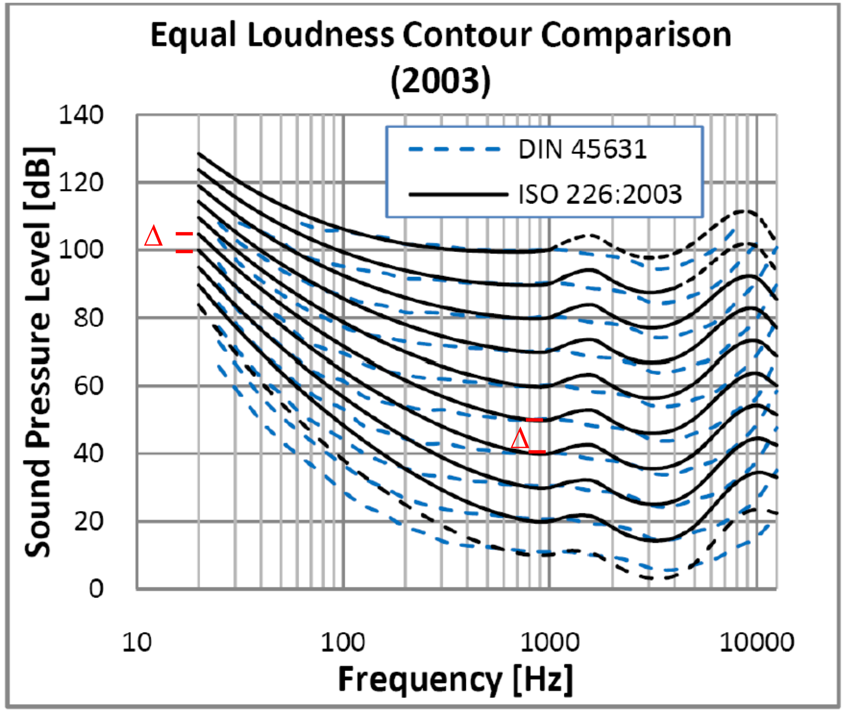
Variation with time, level, frequency, bandwidth, envelope, modulation, etc.

The model must capture the complex function of the auditory system at the periphery and higher levels

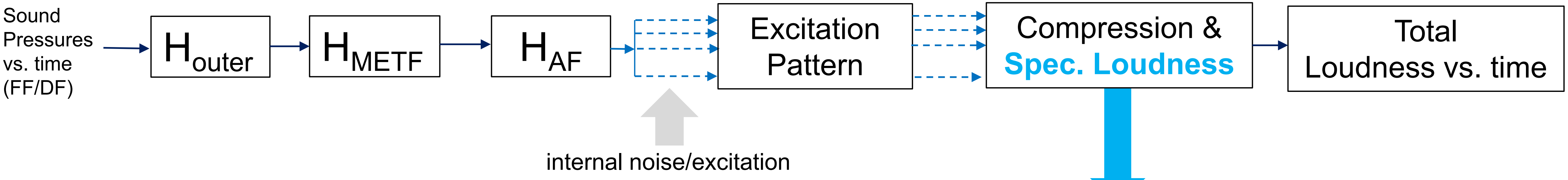
- Nonlinear with both level and frequency
- Frequency masking
- Feedback and focus and tracking of the signal
- Detection below the level of the masker
- Uncertainty
- Binaural hearing attributes

Basic Approach: Reliance on the model of loudness

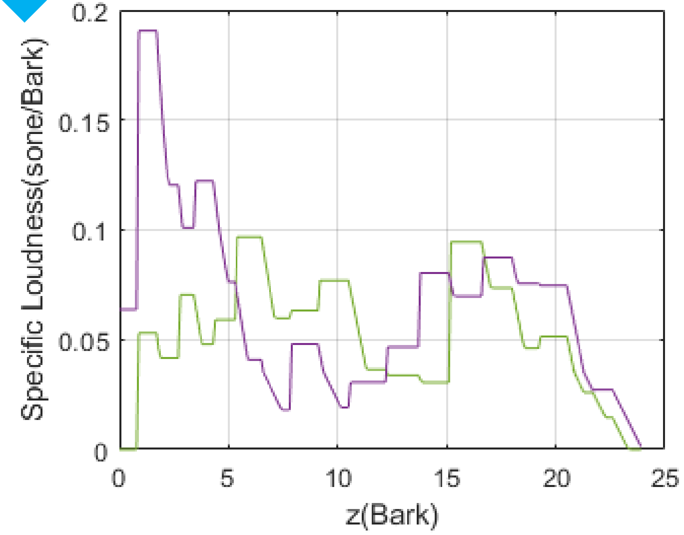
- Nonlinear sensitivity with level and frequency
- Nonlinear to change in level (Δ)
- Absolute threshold of hearing



Specific Loudness



FF/DF – Free/Diffuse Field
H_{outer} – Transfer function to eardrum
H_{METF} – Transfer function, eardrum to differential pressure across Basilar Membrane (BM)
H_{AF} – Auditory Filter (AF) bank



specific loudness at two different time instances for a sample sound

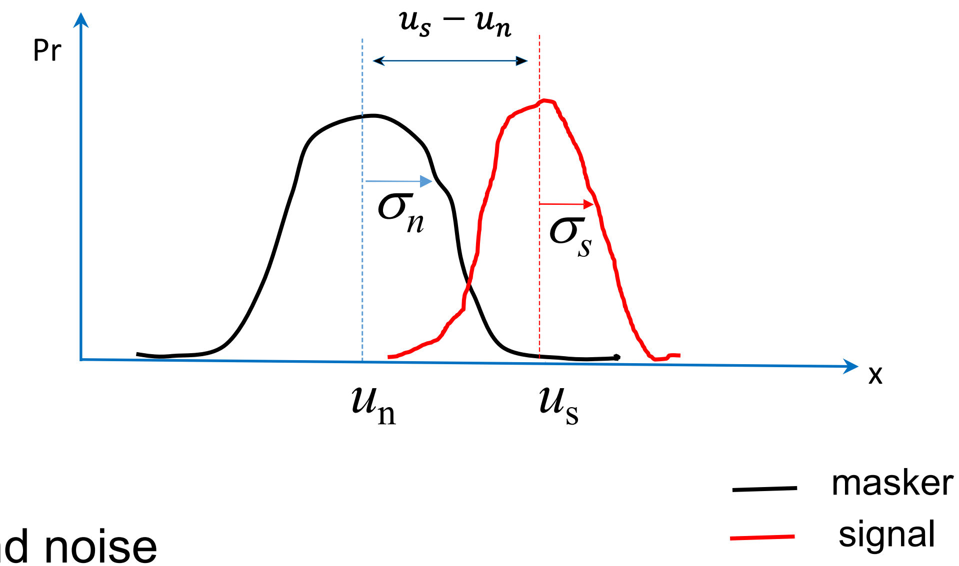
- Specific loudness is the estimate of the strength of sound perceived through different AFs.
- SAP method relies on the specific loudness of the signal and the masker to estimate audibility as it occurs; i.e., with time and frequency.

d-prime or sensitivity

Specific loudness* is assumed to capture "observations" through individual auditory filters. As such, observations arise with a specific probability from either the signal or the masker. According to Signal Detection Theory (SDT), the problem of discrimination (or detection) is a statistical question which relies on testing of statistical hypotheses.

Define sensitivity as:

$$d' = \frac{u_s - u_n}{\sqrt{0.5(\sigma_n^2 + \sigma_s^2)}}$$



with

u_s and u_n - mean of AF response for signal and noise

σ_s and σ_n - standard deviation of response for signal and noise

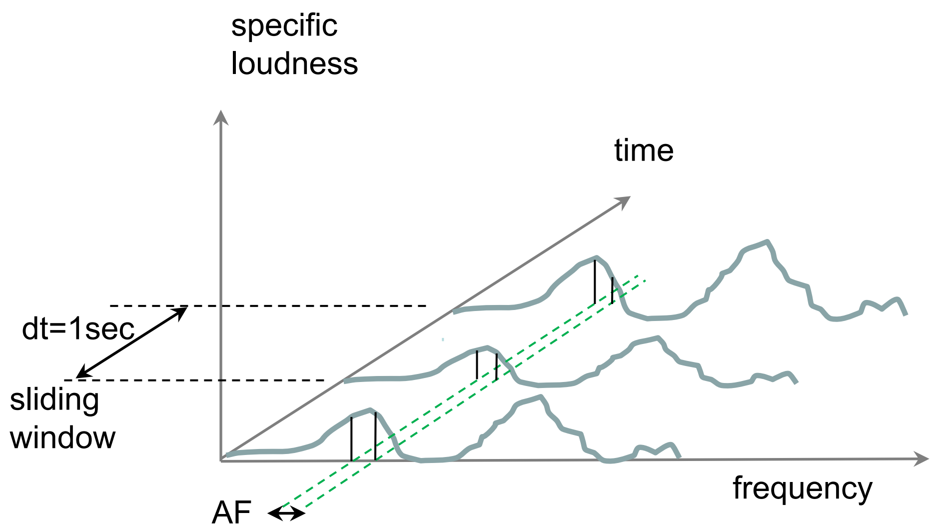
d' measures the mean difference of two distributions normalized to their common standard deviations

* viewed as capturing internal response

SAP Model

1. **Compute specific loudness vs. time for the signal and masker as inputs (at the observer)**
2. **Compute d-prime (sensitivity) within each AF for time span dt:**

$$d'_{t,i} = \frac{\overline{ISPL_{t,i}}}{\sqrt{0.5 * \left(\sigma_{n\,t,i}^2 + \sigma_{s\,t,i}^2\right) + \sigma^2}}$$



$\overline{ISPL_{t,i}}$ mean Instantaneous Specific Partial Loudness (ISPL) for i_{th} AF at time t
σ_s and σ_n - standard deviations for the signal and noise
σ – small correction when signal is minimally present

3. **Compute overall enhanced sensitivity at time t when/if many auditory filters are involved:**

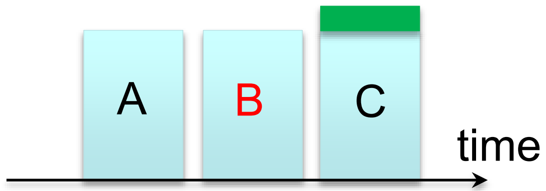
$$d'_t = \left(\sum_{i=1}^{i=n} (d'_{t,i})^2\right)^{\frac{1}{2}}$$ for i = 1, 2, 3, … n=39 AFs   Root Sum of Squares (RSS)

Performance Validation

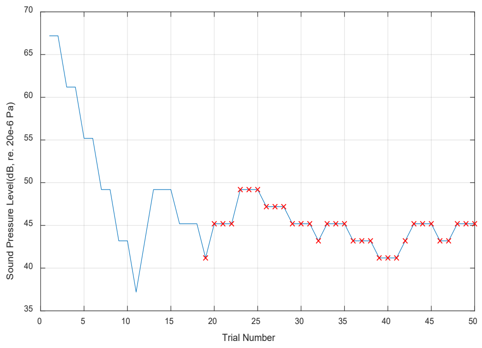
Low frequency tones and a tone complex

Performance Validation (pure tones)

Signal tone frequencies: 55, 120 and 200Hz
Masker: nominal spectral level of 31 dB/Hz and 40-250Hz
Test method: adaptive 3-Alternative Forced Choice
Number of test subjects: 9



3AFC Trial Sequence.
Teal: narrow band noise masker
Green: signal

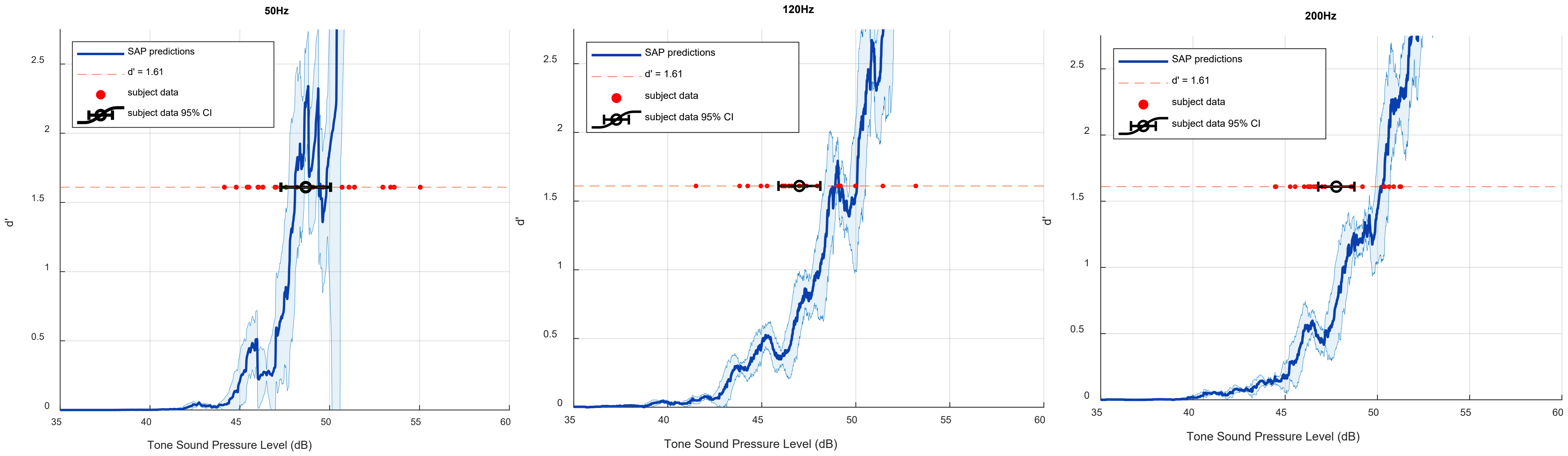


subject response
under adaptive 3AFC test process

P(c) = 79.4% → Hacker and Ratcliff tables → d'=1.61

Rafaelof, M., Christian, A.W., Shepherd, K.P., Rizzi, S.A., and Stephenson, J.H.,
"Audibility of Multiple, Low-Frequency Tonal Signals in Noise," NASA TM-2019-220398, September 2019.

Performance Validation (pure tones)



See common legend

SAP predictions:

Dark blue trace is the mean value of d’

Light blue traces represent the upper and lower %95 limits of the mean = mean -/+ z*SE

assuming normally distributed d’ and computing variance for 1sec long moving window

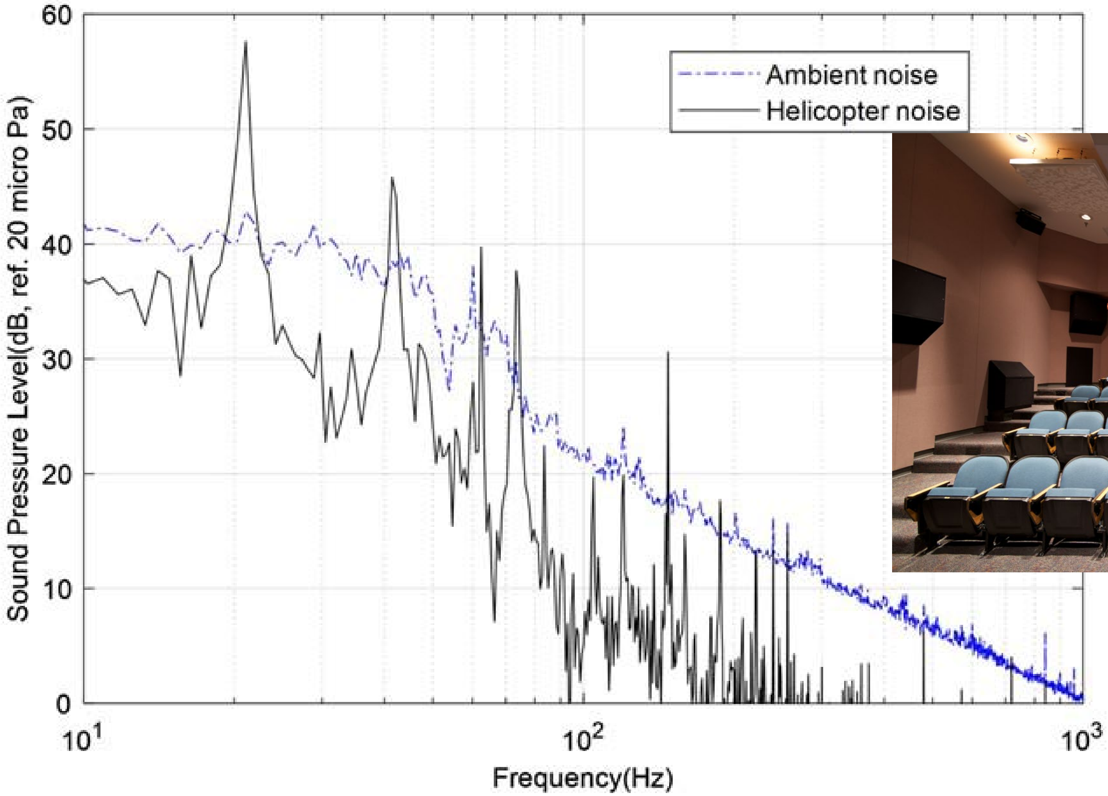
SE = sqrt(variance)

z = 1.96, the 97.5 percental point of the normal distribution

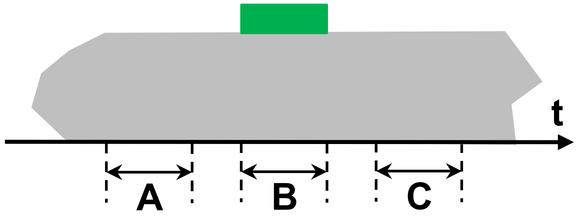
Shaded lighter blue represent the range to the upper and lower %95 limits of the mean

Performance Validation (complex tone)

Signal: fly-in noise by a civilian helicopter
Masker: urban environment ambient noise
Test method: 3-Alternative Forced Choice
Test subjects: 40
Test location: four seats, Exterior Effects Room (EER)



Approach: Source fly-in noise has been sampled 150 times throughout its range; i.e., from completely inaudible to fully audible. This sampling resulted in 150 snippets (800msec in duration) which were presented randomly to subjects over the ambient noise as part of 3AFC test process.

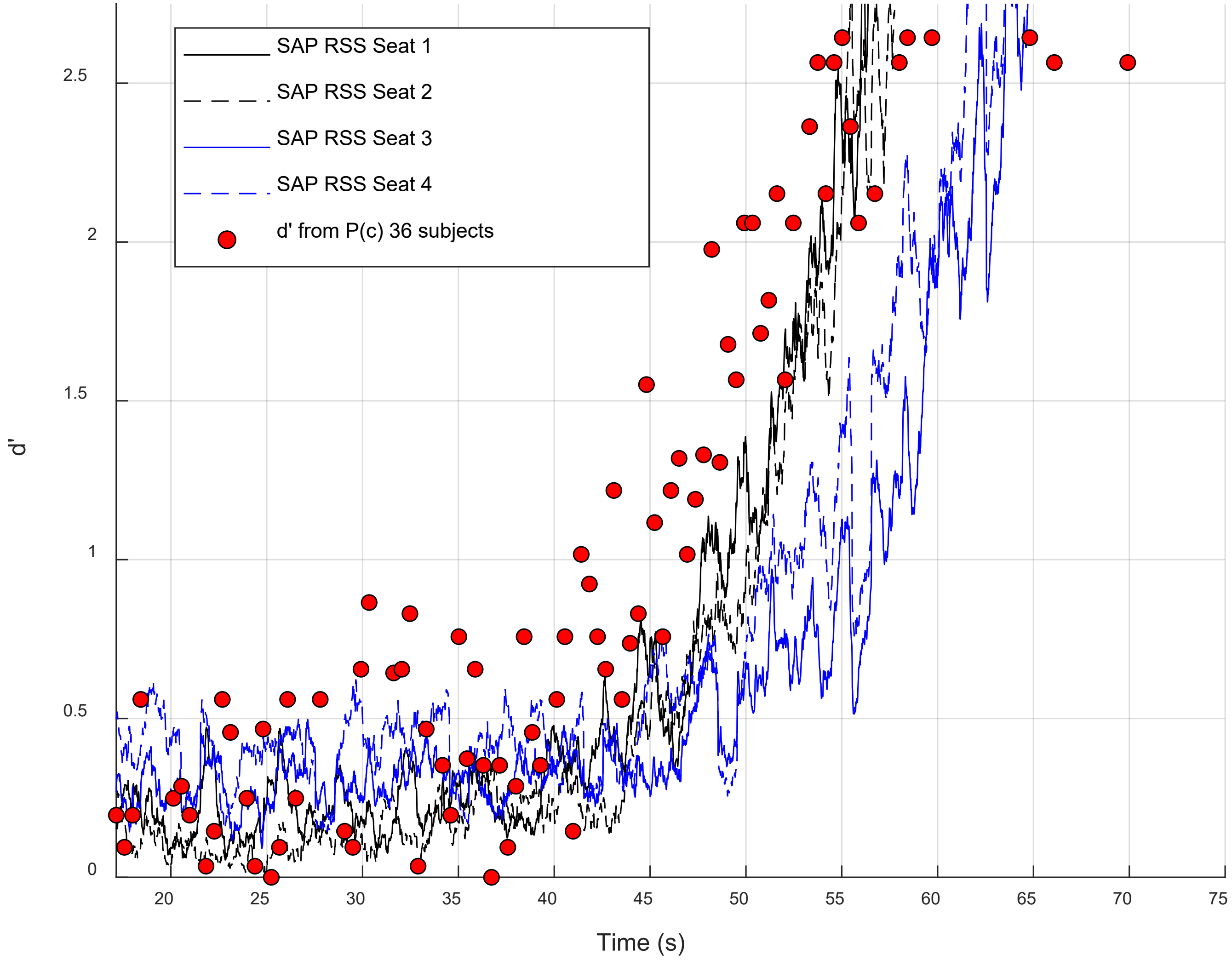


3AFC intervals: green represents the source snippet sound randomly presented over persisting masker (gray).

$$P(c)_{snippet} = \frac{\#\text{ of correct responses}}{\text{total }\#\text{ of responses}}$$

→ Hacker and Ratcliff tables → $d'_{snippet}$

Performance Validation (complex tone)



Mean subject response data for 36 subjects
Front row: seat 1 and 2
Back row: seat 3 and 4

Ongoing Work

1. Accounting for the detection of the signal below the level of the masking noise
2. Binaural hearing
   - Potential improvement due to binaural redundancy vs. current assumption of recruitment
   - The introduction of phase difference (Binaural Masking Level Differences) below 1600Hz
3. Increase in MT due to increased masker bandwidth past the bandwidth of CB
4. Higher frequency signals (2023)
   - Head shadow effect; i.e., different SNRs at ears

1, 2 & 3 : Potential factors explaining earlier detection by subjects

Detection of the signal below the level of the masking noise

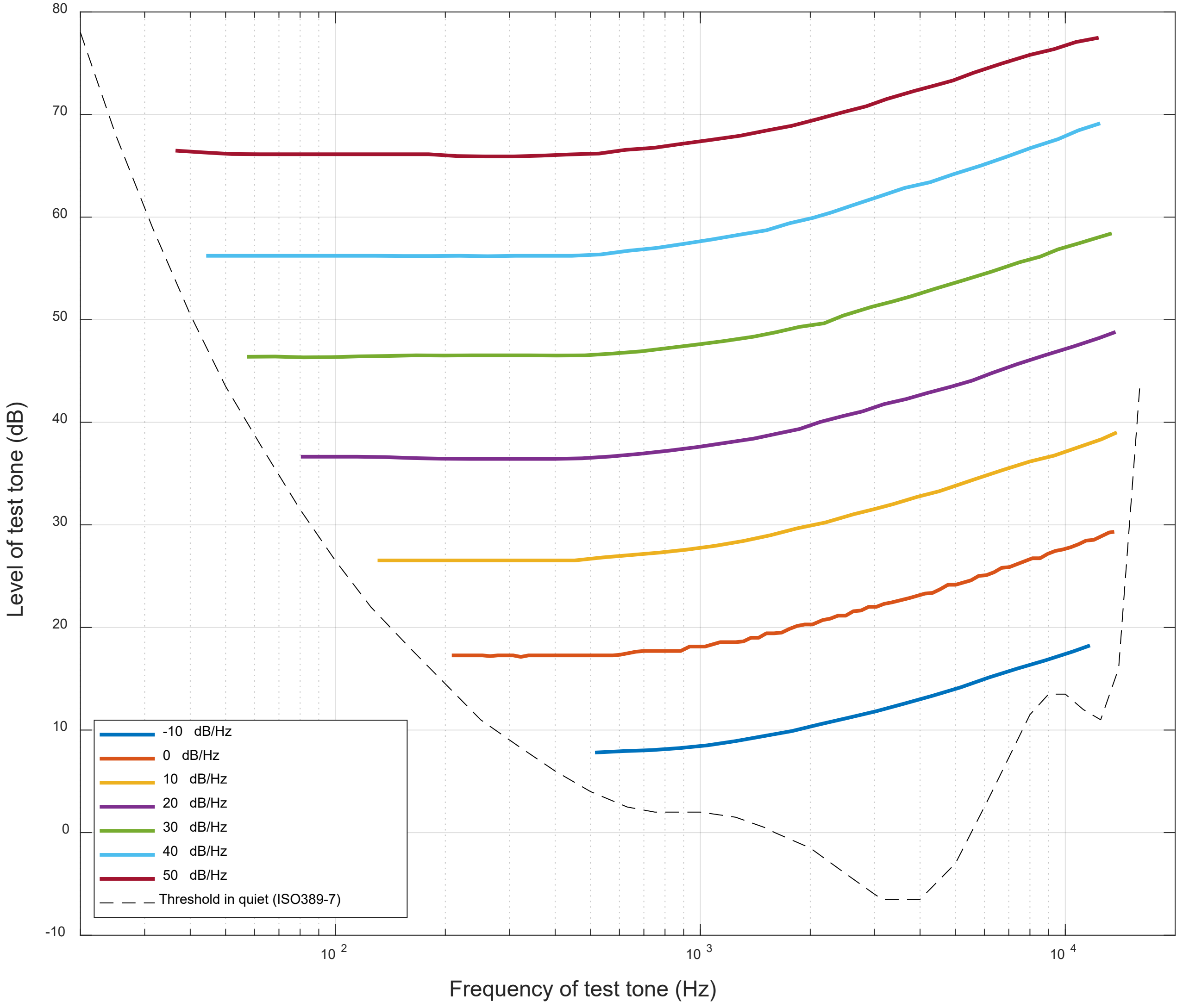
Approach:

- Rely on audibility data in literature for pure tones to check predictions
- Establish a correction to account for the ability for detection below the level of the masking noise

Hearing below the level of the masker

Subject test data*
Masked Threshold (MT) data gathered for pure tones with different frequencies in presence white noise.

Seven curves for seven different noise levels



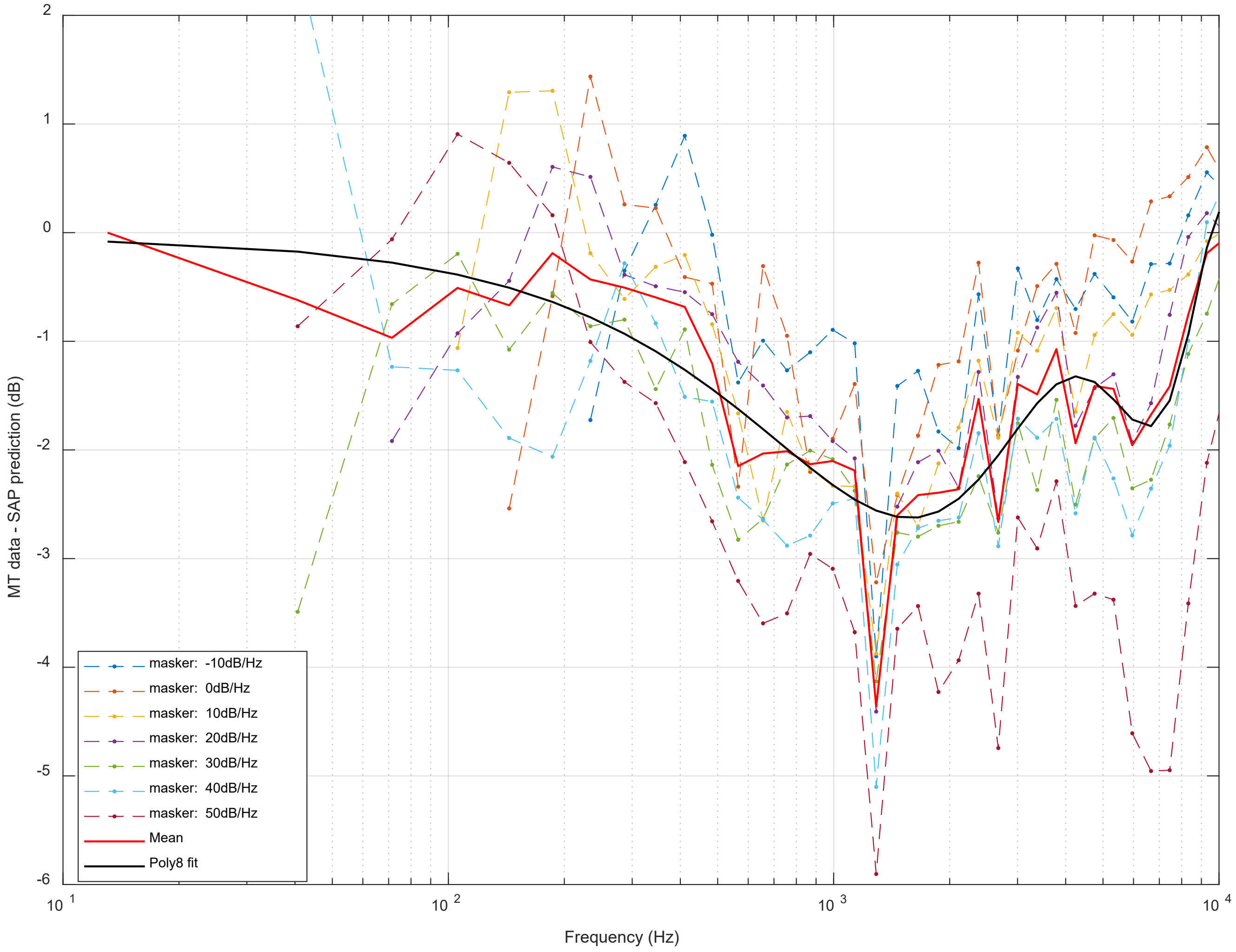
The level of test tone just masked by white noise of different density levels as function of test-tone frequency.

*source: **Figure 4.1** (pp. 62), Psychoacoustics Facts and Models Hugo Fastl and Eberhard Zwicker 3rd edition

Hearing below the level of the masker

Difference between empirical data in Figure 4.1 and predictions by SAP assuming d’ = 1 for many pure tones*.

difference (dB) =
Tone MT – SAP prediction



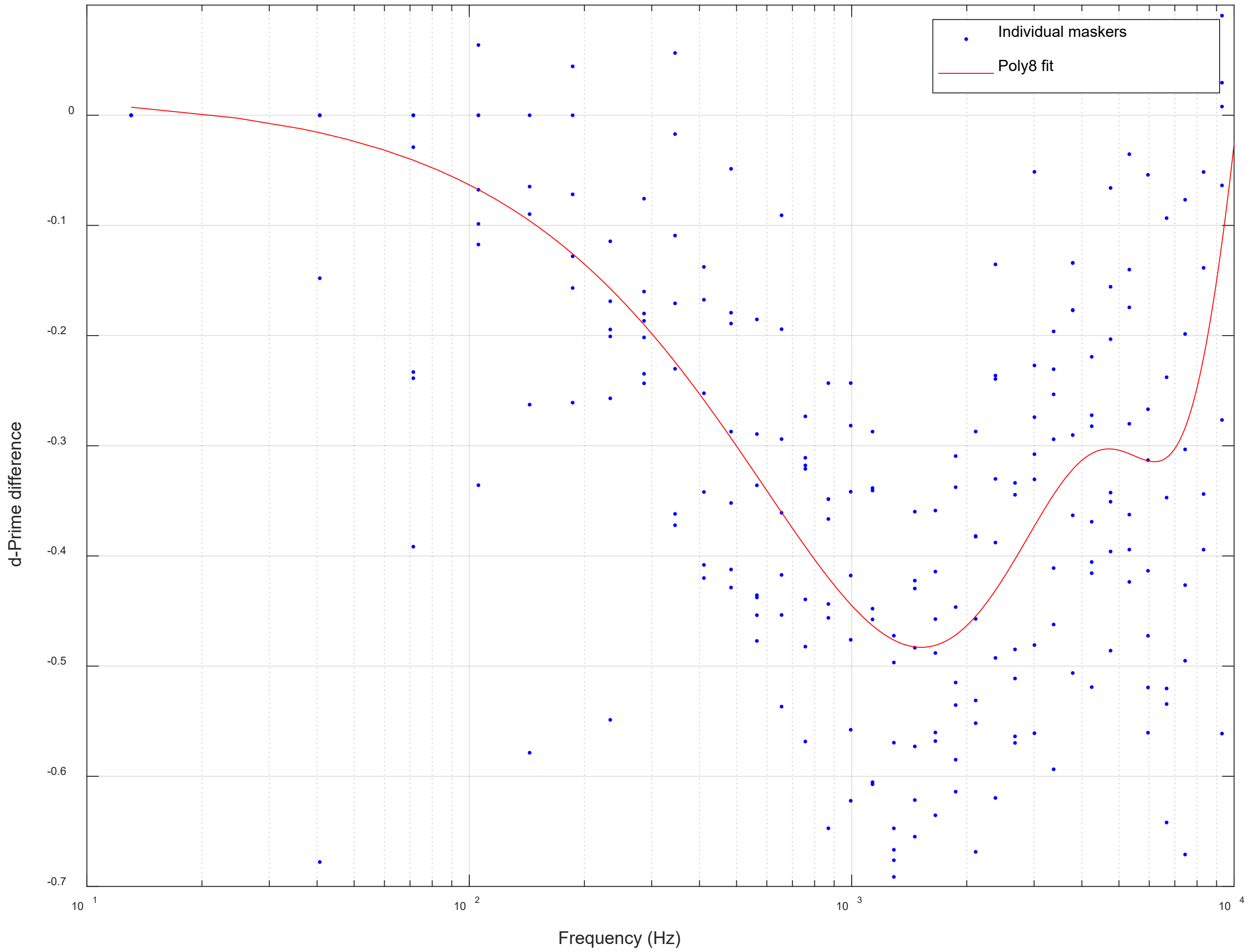
This data points to nonlinear sensitivity of detection with frequency However, mostly, independent of the level.

*39 test tones for each curve

Hearing below the level of the masker

Difference between empirical data in Figure 4.1 and predictions by SAP method expressed as a change in d’.

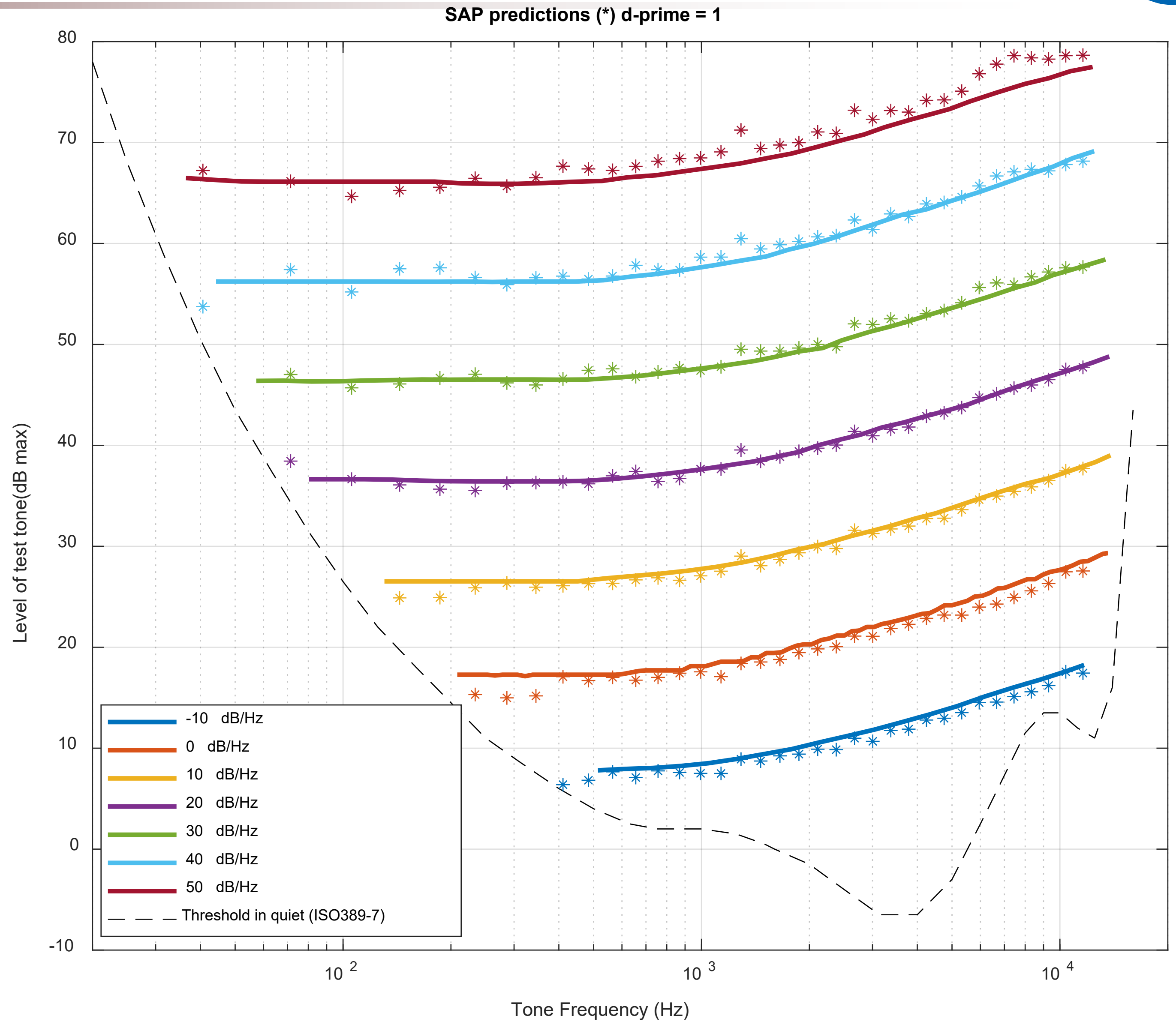
A single d’ correction curve, independent of level, vs. frequency.



Hearing below the level of the masker

SAP Mask Threshold predictions with different ramping-level tones (*) in presence of the same maskers as before.

Pending validation with real data.



*source: Figure 4.1 (pp. 62), Psychoacoustics Facts and Models Hugo Fastl and Eberhard Zwicker 3rd edition

Demonstration

Case	Signal (fly-over sound)	Masker sound
1	Helicopter	NYC ambient
2	SUI drone	UPS truck
3	Distributed Electric Propulsion (DEP) vehicle	box truck

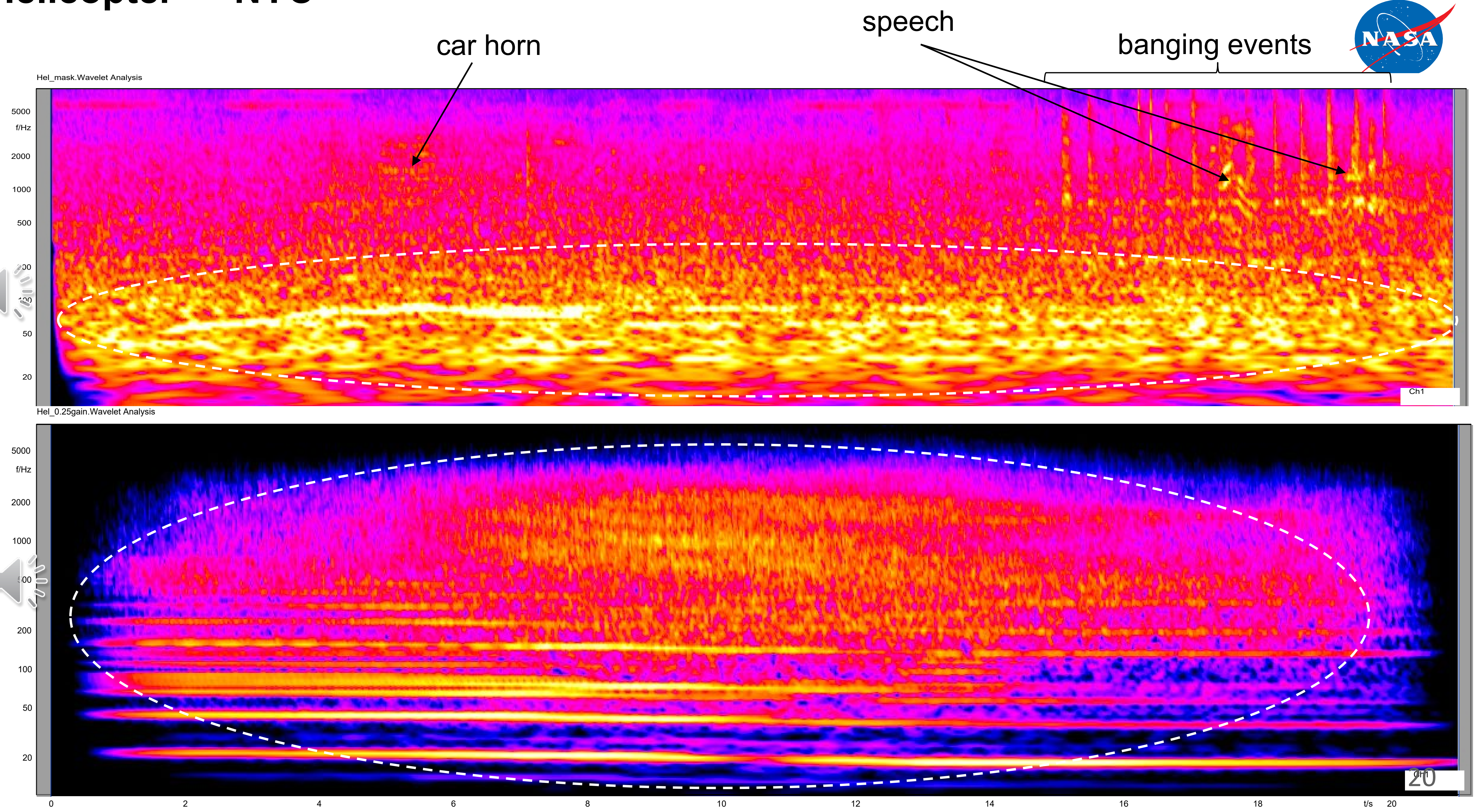
A task for this audience:

Try to consider d’ as a simple noise metric that could gauge intrusiveness or annoyance.
Please provide comments.

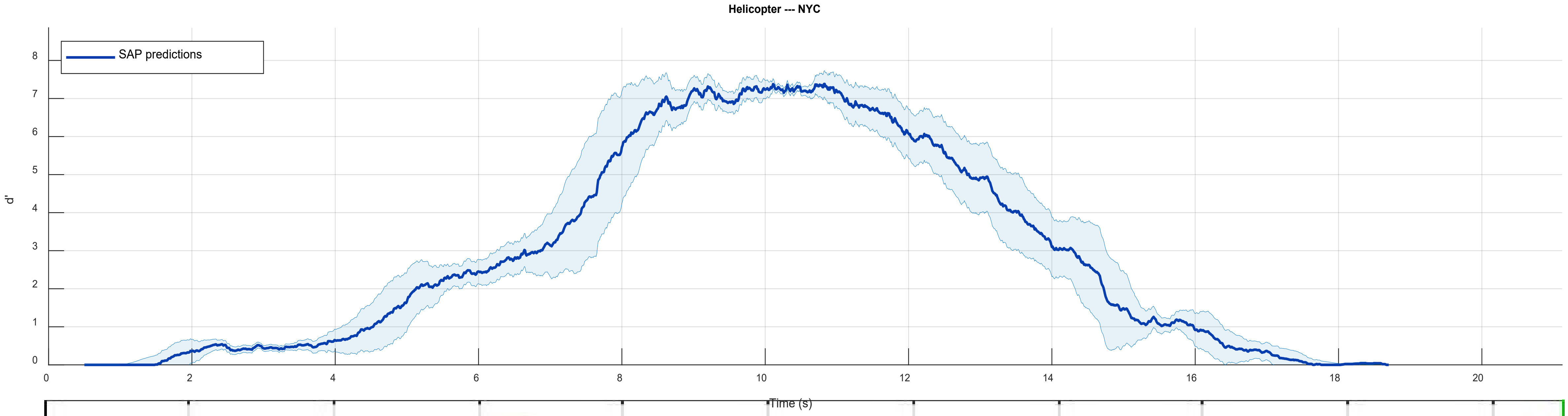
d’ as a noise metric: derived from instantaneous value of d’ vs. time.

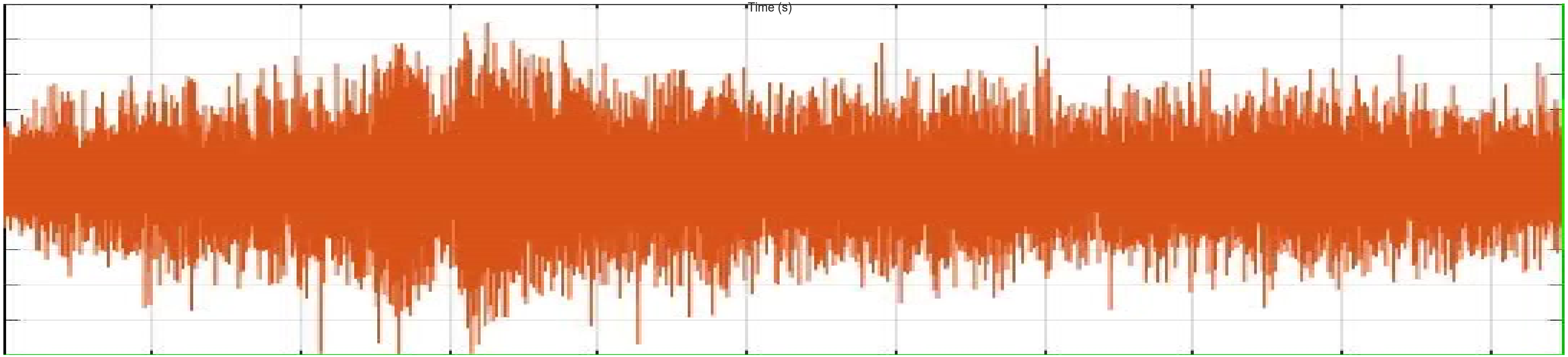
- Peak or rms value of d’ or log(d’) for single events
- Integrated value of d’ or log(d’) with respect to time for cumulative effects

Helicopter --- NYC

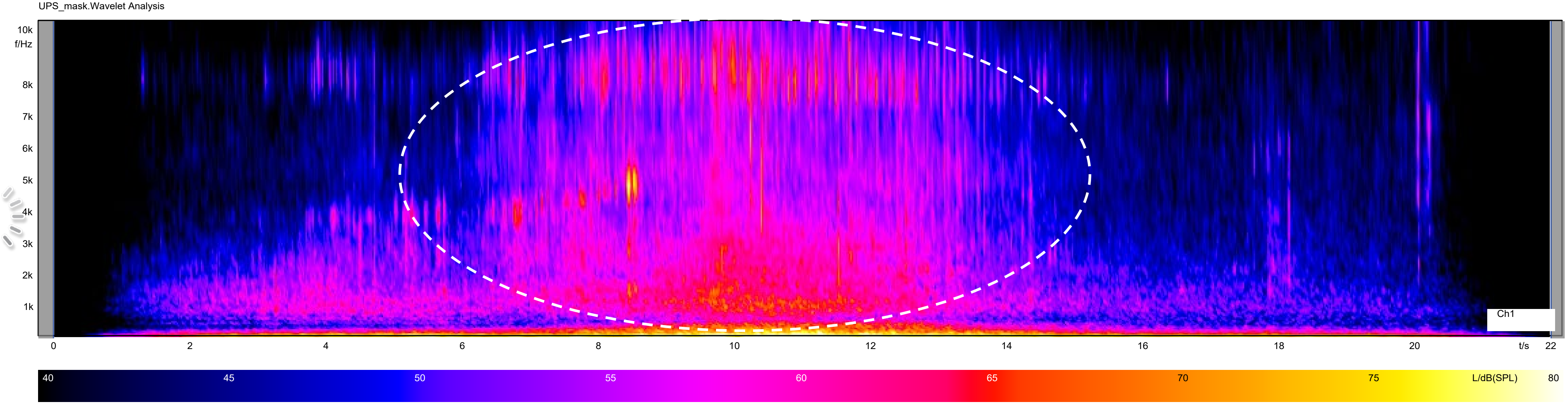


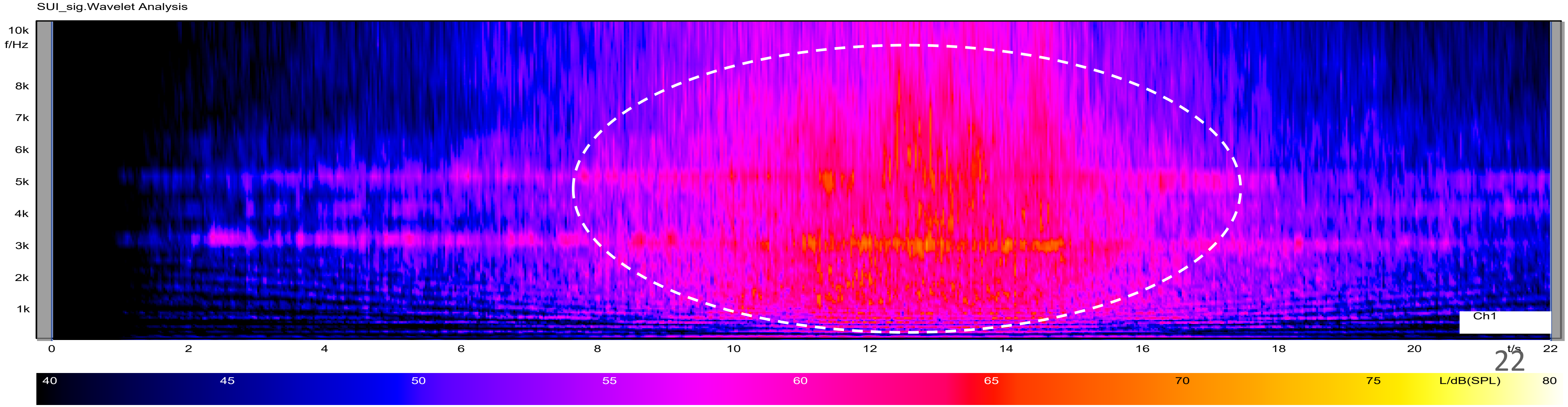
Helicopter --- NYC



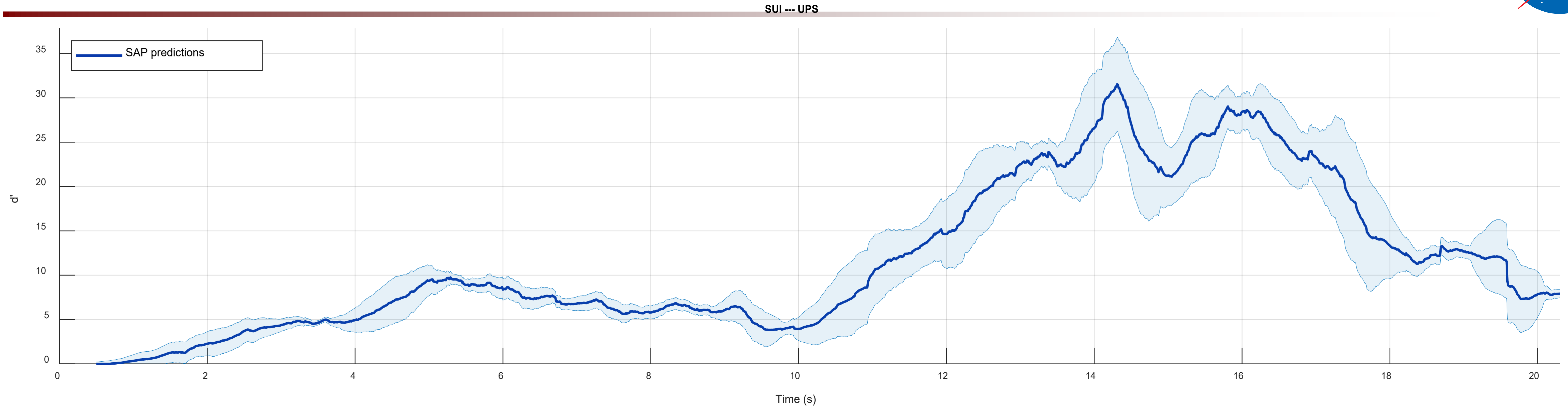


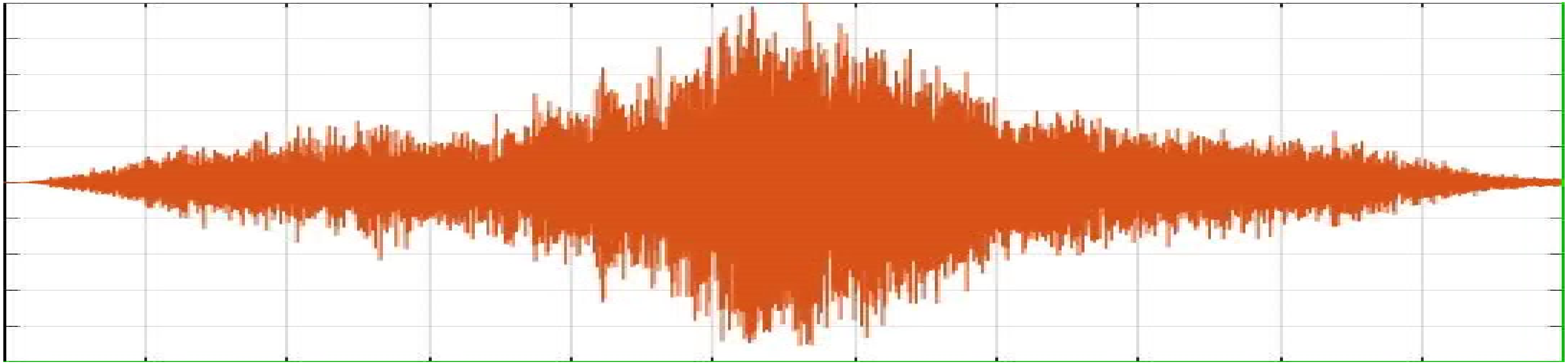
SUI drone --- UPS truck



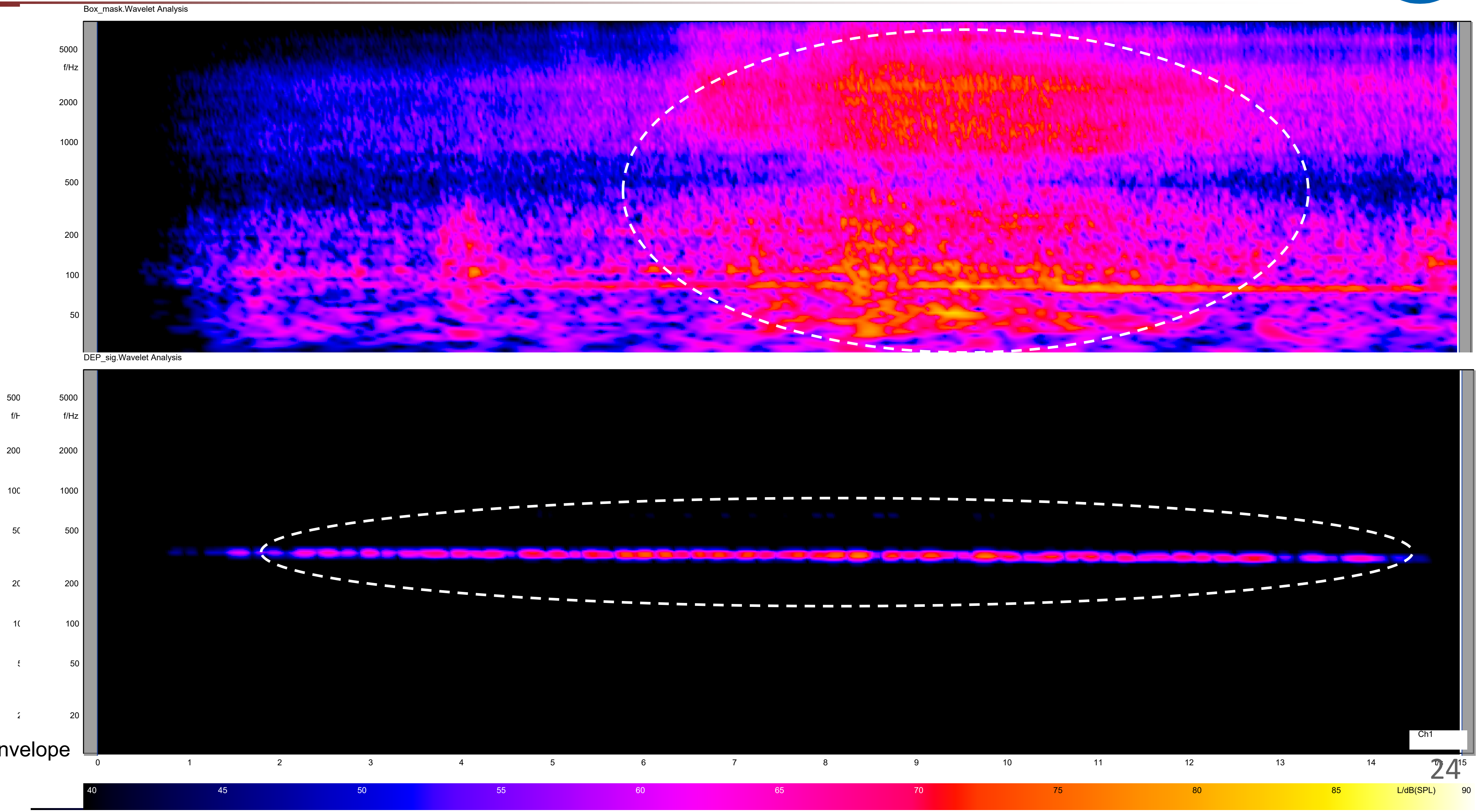


SUI drone --- UPS truck



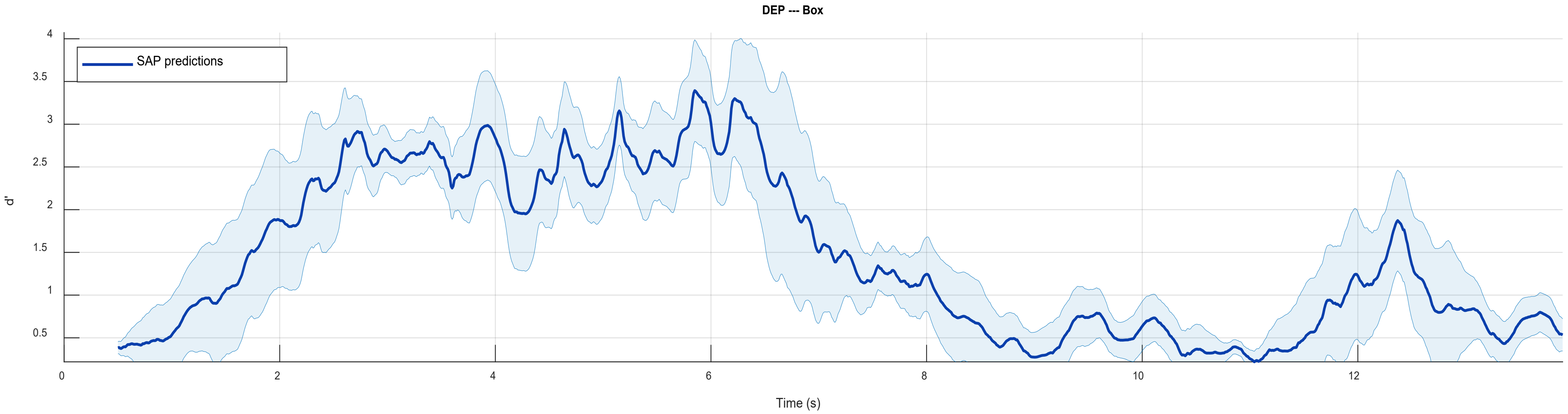


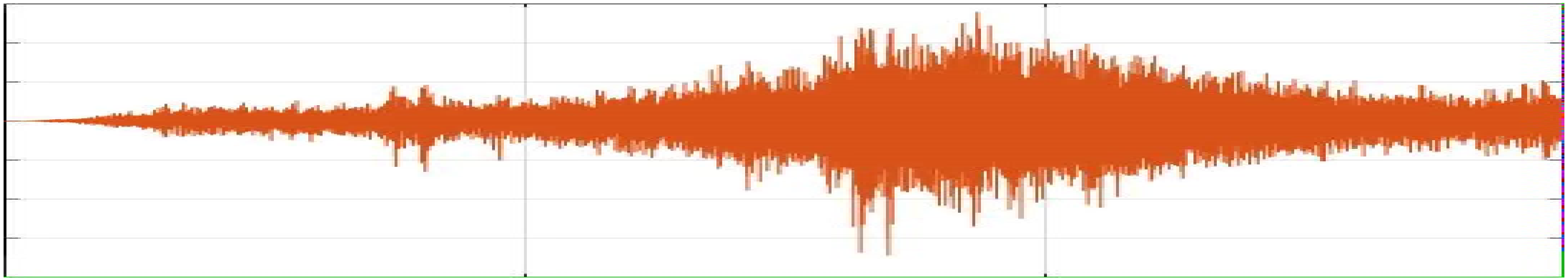
Distributed Electric Propulsion (DEP) --- box truck



Notice envelope

Distributed Electric Propulsion (DEP) --- box truck





Conclusions

- A new modeling approach has been developed based on the hypothesis that audibility is more accurately discerned within individual auditory filters.
- The main attributes of this approach are the prediction of audibility vs. time and frequency, which is essential for achieving high prediction accuracy while enabling the user to identify root cause(s) for audibility. Another attribute of SAP is the ability to predict audibility at a desired probability or sensitivity (d').
- While validation samples were restricted to low-frequency sound, existing data for both discrete tones and complex tones with changing spectrum illustrate the performance of SAP in terms of its prediction accuracy.

Acknowledgment

Thanks to NASA's Revolutionary Vertical Lift Technology (RVLT) project for supporting this work
https://www.nasa.gov/aeroresearch/programs/aavp/rvlt

Thank You