Extending Sottek Hearing Model loudness to estimate partially-masked sound qualities of loudness, tonality, and sharpness

JASA Express Letters

JUNE 16 2026

Extending Sottek Hearing Model loudness to estimate partially-masked sound qualities of loudness, tonality, and sharpness

Michael J. B. Lotinga ; Matt Torjussen; Marc C. Green ; Roland Sottek ; Antonio J. Torija

JASA Express Lett. 6, 063601 (2026)
https://doi.org/10.1121/10.0044195

JASA EXPRESS LETTERS ARTICLE asa.scitation.org/journal/jel

Extending Sottek Hearing Model loudness to estimate partially-masked sound qualities of loudness, tonality, and sharpness

Michael J. B. Lotinga,[1,a)] Matt Torjussen,[1,b)] Marc C. Green,[1,2,c)] Roland Sottek,[3,4,d)] and Antonio J. Torija[1,e)]
[1]*Acoustics Research Centre, Acoustics Innovation Institute, University of Salford, Salford, United Kingdom*
[2]*Arcani Systems, London, United Kingdom*
[3]*Division of Applied Acoustics, Chalmers University of Technology, Gothenburg, Sweden*
[4]*HEAD acoustics, Herzogenrath, Germany*

Abstract: The Sottek Hearing Model provides a means to analyse perceptual sound qualities including loudness, tonality, roughness, and fluctuation strength. An extension to the loudness algorithms is described, which addresses cases in which a source of interest is partially masked by other sounds, such as an encompassing ambient acoustic environment. The proposed extension is validated with objective predictions for unmanned aircraft systems (UAS) and artificial noise, and subjective test results for UAS. The partial loudness extension proposed also enables the calculation of further partial sound qualities, including partial tonality, partial tonal loudness, and partial sharpness. Software implementations of the tools are provided.

[Editor: Andrew Mitchell] https://doi.org/10.1121/10.0044195

Received: 2 April 2026 **Accepted:** 4 June 2026 **Published Online:** 16 June 2026

1. Introduction

Psychoacoustic models of human perception enable the assessment of sound qualities and a rich description of the aural impressions that may be expected for a given acoustic signal. The Sottek Hearing Model (SHM) offers a holistic psychoacoustic sound quality framework, which includes sound quality metrics (SQMs) for perceptual attributes of loudness, sharpness, roughness, tonality, and fluctuation strength (Ecma International, 2025; Sottek, 1993).

When the focus is on a specific sound source operating within a real environment, there may be a particular interest in understanding how perception of the source sound qualities could be influenced by other sources in the acoustic environment. For example, complex perceptual sound masking effects can reduce the audibility of the source, leading to it appearing as less prominent in the overall soundscape. A psychoacoustic model for estimating the loudness of partially-masked sources was developed by Moore *et al.* (1997) and Glasberg and Moore (2005), and this description of "partial loudness" has been applied to analyse the potential for ambient acoustic environments to mask the sound immissions of conventional transportation sources (Lee *et al.*, 2021), rotorcraft, and advanced air mobility (AAM) vehicles (Rafaelof and Wendling, 2021; Schlittenlacher and Moore, 2021).

In this study, an extension of the partial loudness concept is conceived, which proposes that other partially-masked sound quality metrics (P-SQMs), beyond loudness, can also be evaluated, providing a more comprehensive psychoacoustic description of the perception of an acoustic signal within other masking sounds. The SHM provides a platform for realising this idea, which is based on using a form of partial loudness to analyse other sound qualities, including tonality, tonal loudness, and sharpness. To achieve this, a partial loudness based on the SHM has been developed and implemented. Prediction results from this P-SQM are compared with those obtained using an implementation of the time-varying adaptation of the Moore–Glasberg partial loudness model for a particular application case of interest: the sounds of unmanned aircraft systems (UAS) operating within ambient acoustic environments. Subjective test data from a psychoacoustic experiment have also been used to validate the SHM-based partial loudness (Sec. 4). The straightforward manner in which the P-SQMs for partial tonality, partial tonal loudness, and partial sharpness proceed from the SHM-based partial loudness is then described (Sec. 5), alongside a metric for (partial) "tonal sharpness," defined here as the sharpness of the (partial) tonal loudness.

[a)]Corresponding author: m.j.lotinga@edu.salford.ac.uk
[b)]Email: m.c.torjussen@edu.salford.ac.uk
[c)]Email: m.c.green@salford.ac.uk
[d)]Email: roland.sottek@chalmers.se
[e)]Email: a.j.torijamartinez@salford.ac.uk

The notation employed in this work follows the formulaic definitions of the standardised SHM SQMs (Ecma International, 2025), with full details of the implemented algorithms and parametric values provided by Lotinga (2026). Software implementations for all of the SHM-based P-SQMs presented have been developed in MATLAB, licensed for open-access use and made publicly available (Lotinga, 2026); parametric inputs to all formulations presented can be found in the software code.

2. SHM loudness overview

The loudness SQM in the SHM is standardised in ECMA-418-2 (Ecma International, 2025), established from the theoretical consideration and empirical validation reported by Sottek (1993, 2016). Briefly, the acoustic pressure signal input is filtered to emulate the effect of the human outer and middle ear, followed by a further bank of critical band (z) auditory filters, representing the basilar membrane spectral selectivity. The filtered z-dependent signals are segmented into overlapping time blocks (l), transformed using a non-linear compressive function, and adjusted for the loudness threshold in quiet (LTQ) to obtain the "specific basis loudness" N'_{basis}. Autocorrelation analysis is applied to identify and separate the tonal and broadband components, which are individually processed to reduce signal processing noise, resulting in separated estimates of tonal and broadband partial loudness; these components are combined using a weighted sum to obtain the psychoacoustic specific loudness $N'(l,z)$ for the input pressure signal. Compared with the loudness model of Moore *et al.* (1997), the SHM loudness covers a wider frequency range and yields predictions in closer agreement with conventional empirical equal-loudness contours (Sottek *et al.*, 2023). Details of the SHM loudness SQM definition can be found in the standard (Ecma International, 2025), and a verified, open-source implementation has also been presented by Lotinga *et al.* (2025b) and Lotinga (2026).

3. Partial loudness extension to SHM loudness

The defining principles behind this partial loudness extension to the SHM are as follows. The partially-masked loudness of the target signal $N_{\text{part,target}}$ should adhere to these boundary conditions:

1. When the target signal excitation is well below the masking noise excitation, $N_{\text{part,target}}$ approaches 0.
2. When either of the following two conditions hold, $N_{\text{part,target}}$ approaches the value it would have if the masker was not there (N_{target}), i.e., when
   (a) either the target signal excitation is well above the masker, or
   (b) the masking noise is below excitation threshold in quiet (ETQ),
   partial loudness approaches its unmasked value.
3. When the target signal and masking noise have equal excitation, $N_{\text{part,target}}$ approaches what it would be if the target signal excitation was equal to ETQ in the absence of the masker.

The approach taken starts with calculation of the quantity $\tilde{N}'(l,z)$ described in ECMA-418-2 as the "*specific loudness of the signal without consideration of the threshold in quiet*" (Ecma International, 2025) separately for both the target signal and masking noise signals. Then, a simplified function is applied to approximately invert the non-linear compressive loudness transformation using band-dependent parameters, yielding quantities resembling auditory excitation patterns. The difference between masker and target in this "inverse-compressed" domain is used to apply a dynamic, frequency-dependent weighting to the target loudness $\tilde{N}'_{\text{target}}$ in response to the masker loudness $\tilde{N}'_{\text{masker}}$.

To develop the inverse-compressive function, use is made of a detection efficiency factor K, presented by Moore *et al.* (1997) as an empirical quantity representing the ratio of target signal excitation to masking noise excitation "*at the output of the auditory filter required for threshold at high masker levels.*" Here, the value of K is approximated using a 4th-order polynomial expression over the base-10 logarithm of the critical band centre-frequencies $F(z)$, after Ward (2019). A band-dependent masking function compression parameter $\gamma_c(z)$ is then defined using limiting constants q_n (see the supplementary material for parameter values and a signal processing flow diagram),

$$\gamma_c = q_1 + (q_2 - q_1)\frac{K}{1+K}. \tag{1}$$

The inverse-compressed loudness quantities $\nu(l,z)$ are calculated for both target and masker using γ_c as follows:

$$\nu = \tilde{N}'^{(1/\gamma_c)}. \tag{2}$$

The subtraction of masker from target is then performed using these inverse-compressed quantities, normalised (adding $\epsilon = 10^{-12}$ to avoid any undefined ratios), ensured non-negative, and exponentiated using a damping parameter, γ_d, defined in the same way as in Eq. (1), using limiting constants q_3, q_4; this γ_d controls the smoothness in the response of Eq. (3):

$$\tilde{\nu}_{\text{target}} = \max\left(0, \frac{\nu_{\text{target}} - 0.5\nu_{\text{masker}}}{\nu_{\text{target}} + \nu_{\text{masker}} + \epsilon}\right)^{\gamma_d}. \tag{3}$$

The result of Eq. (3) is then applied as a scaling to ν_{target} (reducing the effective excitation in relation to the masker), with the scaled product "recompressed" yielding an estimate of the specific partial loudness of the target, prior to consideration of the LTQ,

$$\tilde{N}'_{\text{part,target}} = \left(\nu_{\text{target}}\tilde{\nu}_{\text{target}}\right)^{\gamma_c}. \tag{4}$$

Commensurate with the SHM, $\tilde{N}'_{\text{part,target}}$ is then compensated by the z-dependent LTQ to yield an estimate of the specific "partial basis loudness" for the target signal

$$N'_{\text{part,basis,target}} = \max\left(0, \tilde{N}'_{\text{part,target}} - \text{LTQ}\right). \tag{5}$$

This quantity, $N'_{\text{part,basis,target}}$, proceeds in place of N'_{basis} through the remaining processing steps in the same manner as defined in ECMA-418-2 to arrive at the main output, the specific partial loudness for the target $N'_{\text{part,target}}(l, z)$. This $N'_{\text{part,target}}$ in turn also has the standard spectral and temporal aggregations applied to derive time-dependent partial loudness $N_{\text{part,target}}(l)$ and overall (power-time-averaged) partial loudness $N_{\text{part,target}}$. Similarly, in the case of two-channel binaural input signals (for both target and masker), it is assumed that the same binaural quadratic averaging approach defined in ECMA-418-2 can also be applied to $N'_{\text{part,target}}$ to estimate the combined binaural partial loudness.

4. Validation

Two forms of comparison are examined to demonstrate the validity of the developed algorithms: first, Sec. 4.1 presents predictions of SHM-based partial loudness compared with equivalent objective results obtained using an implementation of the existing Moore–Glasberg partial loudness model (Glasberg and Moore, 2005; Moore *et al.*, 1997; Ward, 2017, 2019); second, predictions made using both partial loudness models are compared with subjective noticeability results obtained from a laboratory experiment (Lotinga *et al.*, 2025a), as described in Sec. 4.2.

4.1 Comparison with objective predictions

The sounds used for the comparison with Moore–Glasberg partial loudness comprise binaural recordings of auralised experimental stimuli featuring UAS flight operations within ambient acoustic environments. As detailed by Lotinga *et al.* (2025a), the experiment included three different UAS types and two ambient environments; a selection of results is presented here, and further examples are provided in the supplementary material. The comparison results for time-dependent partial loudness illustrate a good agreement between the predictions made using each model with similar patterns observed, albeit with different absolute loudness scaling (Fig. 1); this difference is expected, since the models produce differing estimates of absolute loudness (Sottek *et al.*, 2023). The two examples presented demonstrate the behaviour of the partial masking: within the relatively steady street ambient environment [Fig. 1(a)], the UAS overflight is predicted to be inaudible (zero partial loudness) for parts of the gradual rise/fall periods of the loudness envelope, whereas the unmasked loudness would be non-zero for almost the entirety of the recording duration. The peak of the overflight event (at 12.5-s) remains higher in (partial) loudness relative to the (absolute) loudness of the ambient but is somewhat reduced by the masking sound. In the more-variable park ambient environment [Fig. 1(b)], brief, transient sound events can be observed as producing sudden masking of the UAS landing event (noticeably at 10-s and 17-s), momentarily reducing the partial loudness—these short events correspond with instances of birdsong rising above the steadier sources in the environment.

The specific partial loudness results (Fig. 2) further illustrate the similarity in behaviour of the models across the auditory frequency range. (When viewing this graphical representation, it should be noted that the visible patterns are sensitive to the selection of the scale upper limit. Differences in time and spectral resolutions between the models also influence the relative magnitudes observed.) For the overflight event [Fig. 2(a)], the partial loudness peak can be seen to be focussed in the 1–6 kHz range for both models. The intermittent, transient birdsong masking effect on the landing event is again clearly visible in the park environment, in the 1.6–4 kHz range [Fig. 2(b)].

Comparing time-aggregated partial loudness predictions for the full sample of recordings again demonstrates that there is good agreement between the results for each model (Fig. 3), despite the differences in their original formulations and corresponding assumptions (Ecma International, 2025; Moore *et al.*, 1997).

A comparison of objective predictions made using artificial white noise signals with varying signal-to-noise ratios can be found in the supplementary material, which also indicates comparable behaviour between the proposed SHM extension and the Moore *et al.* (1997) partial loudness model, and adherence to the boundary conditions defined in Sec. 3.

4.2 Comparison with subjective evaluations

The approach taken in the experiment used to obtain data for the comparisons presented in the present work has been reported in detail by Lotinga *et al.* (2025a). Briefly, the subjective data presented here relates to the judgements of 30 participants, who were exposed to the stimuli outlined in Sec. 4.1, presented over an ambisonic loudspeaker array. The

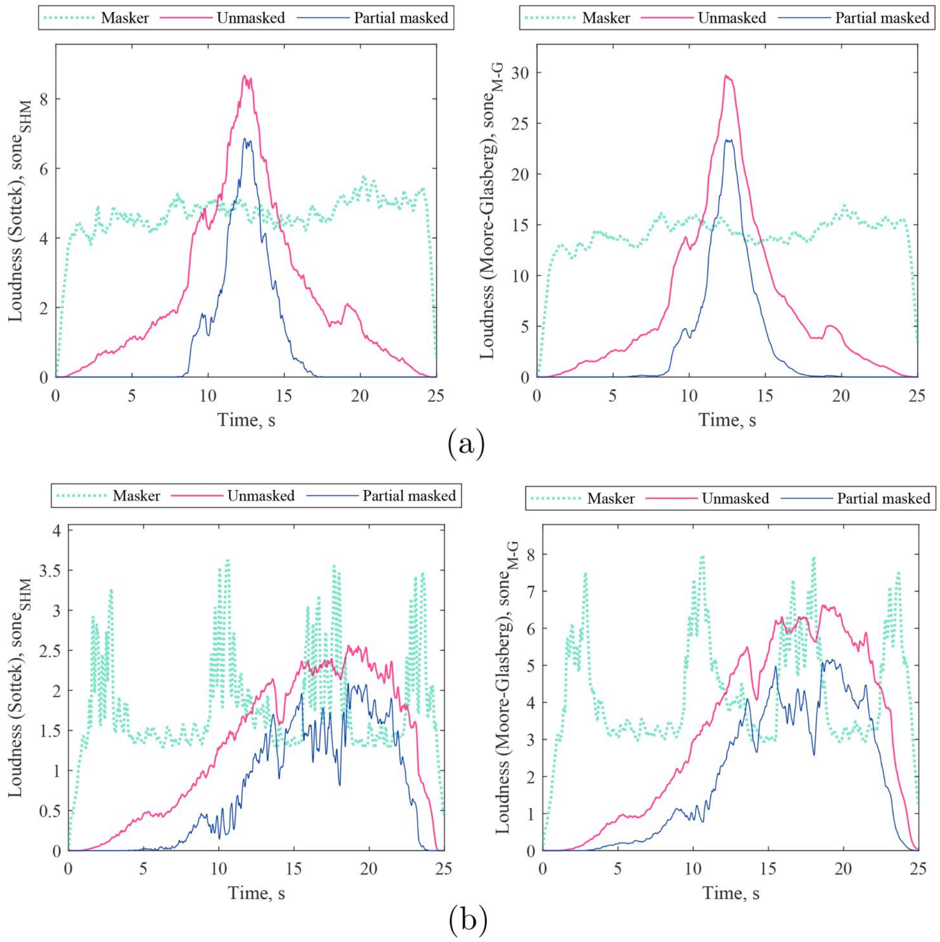

Fig. 1. Comparison of binaurally-combined partial loudness predictions; UAS data from Lotinga *et al.* (2025a): top, UAS overflight with street ambient; bottom, UAS landing with park ambient; solid lines, target UAS sound; dotted lines, ambient masker.

subjective results considered here were derived from participant typological identification of sounds they had "noticed" during each of the acoustic scene stimuli; these judgements were analysed to yield estimates of the percentage of the participants who were considered to have noticed a UAS event within each scene. Some of the scenes had a relatively high loudness of the UAS relative to ambient sound, while others had UAS events with lower relative loudness, which were expected to be partially (or completely) masked by the ambient sound. In this way, the noticing of the UAS event is expected to be directly associated with its relative audibility within each ambient environment.

The comparison of these subjective data with the partial loudness predictions demonstrates very similar behaviour for each model, and both indicate similarly strong predictions of the "UAS noticed" outcome [Figs. 4(a) and 4(b)]. Also shown is a comparison of the subjective data with a third metric, "detectability" [Fig. 4(c)], which provides a means of estimating the expected chance that a masked sound would be aurally detected by a listener on a statistical basis (Fidell *et al.*, 1974; Rizzi *et al.*, 2026). This metric is noteworthy here, as it is similar in application use-case to partial loudness, but operates on considerably lower-resolution data; an implementation of the aural detectability metric is also made available within the same open-source repository as the proposed SHM-based partial loudness (Lotinga, 2026).

5. Partially-masked sound qualities

The results presented in Sec. 4 suggest that the proposed partial loudness based on the SHM produces comparable results to the existing Moore–Glasberg model. A considerable advantage of the new model is that, combined with the holistic approach in the SHM, this development enables further P-SQMs to be defined, building from the (partial) basis loudness. The provided software implementations demonstrate initial examples of these (Lotinga, 2026), which presently include partial tonality, partial tonal loudness, and partial sharpness. Dropping the "target" descriptor, the partial (specific) tonality T'_{part} and partial (specific) tonal loudness $N'_{T,\mathrm{part}}$ are obtained in the standard way as defined in ECMA-418-2, but substituting $N'_{\mathrm{part,basis}}$ in place of N'_{basis}. Partial sharpness S_{part} is derived from N'_{part} using the loudness substitution

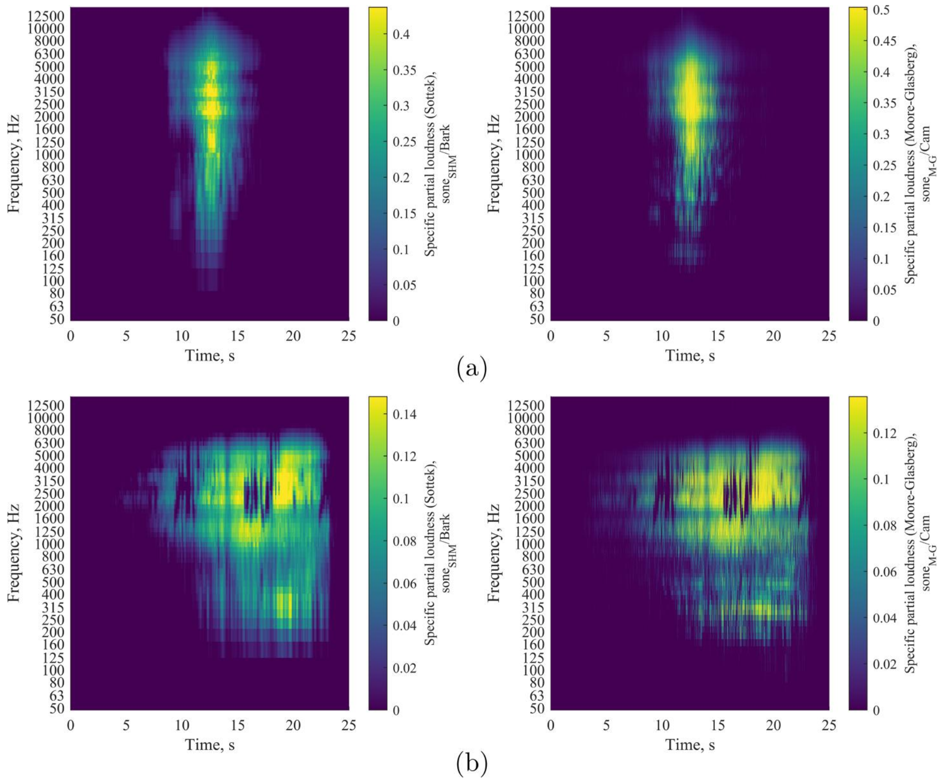

Fig. 2. Comparison of single-channel (left ear) specific partial loudness predictions (corresponding with Fig. 1); UAS and ambient audio data from Lotinga *et al.* (2025a): top, UAS overflight with street ambient; bottom, UAS landing with park ambient.

approach proposed by Swift and Gee (2017).[1] Combining these concepts together, metrics for "tonal sharpness" and "partial tonal sharpness" have also been developed (Lotinga, 2026). The tonal sharpness metric is derived by evaluating the sharpness of the specific tonal loudness N'_T and then calibrating the result to yield 1 acum for a reference signal, which, for consistency with the reference signals for sharpness (Fastl and Zwicker, 2007) and tonality

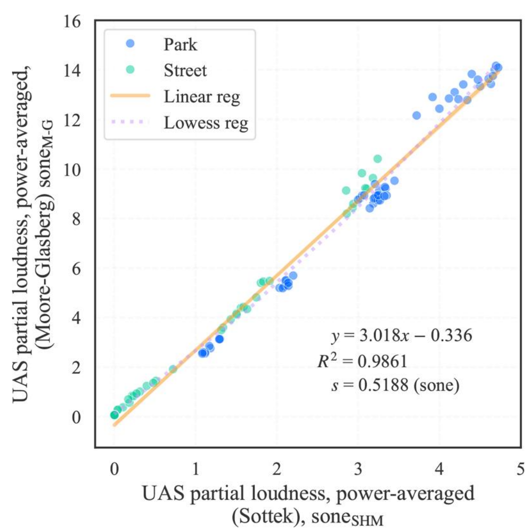

Fig. 3. Comparison of overall binaurally-combined, time-aggregated partial loudness predictions; UAS and ambient audio data from Lotinga *et al.* (2025a); linear and locally weighted scatterplot smoothing regression fits indicating central tendency; corresponding equation, coefficient of determination and residual standard error for linear fit displayed as black inset text.

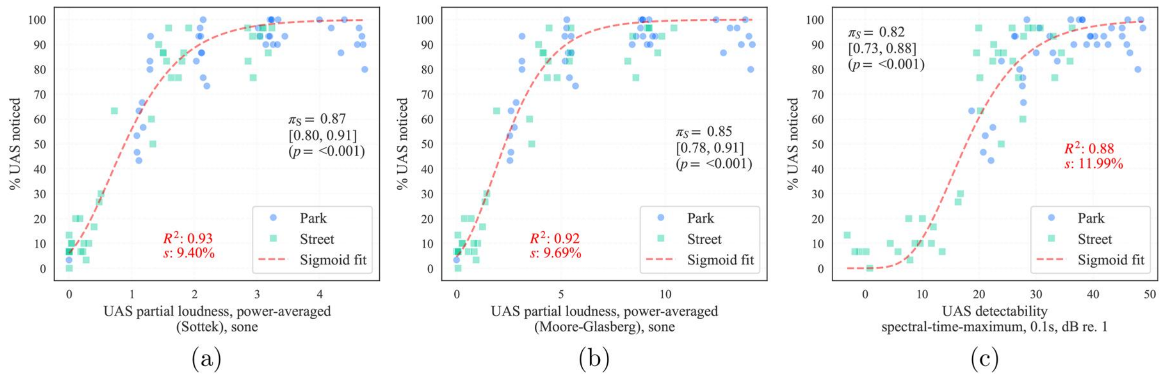

Fig. 4. Comparison of predictions with subjective noticeability data (% values from 30 participants for 80 stimuli) for UAS; markers, data points from Lotinga *et al.* (2025a); Shepherd's π_S robust correlation coefficients with 95% confidence intervals and corresponding p-values displayed as black inset text; red dashed lines, sigmoidal least-squares regression fit indicating central tendencies; coefficients of determination and residual standard errors for sigmoid fits displayed as red inset text.

(Ecma International, 2025), is chosen as a 1 kHz sinusoid in free-field plane wave frontal incidence at 60 dB L_p. The short step to partial tonal sharpness is taken by evaluating sharpness using $N'_{T,\text{part}}$ in place of N'_T.

We defer perceptual comparison and interrogation of these proposed P-SQMs to further research (example calculation results are provided in the supplementary material). The objective of this work is merely to introduce the principles and elucidate the processing approach, which is built from the proposed partial loudness extension to the SHM loudness. The extension itself represents an initial approach, which has been developed by approximating the behaviour of the Moore *et al.* (1997) model using relatively simple algorithms; a more rigorously underpinned approach based on theoretical considerations within the SHM framework could be established in the future. Further refinement via perceptual testing could also enable optimisation of parameters used in the SHM-based partial loudness model, which has not yet been undertaken.

The full utility of these P-SQMs needs to be tested and investigated further; however, we envisage they will be very helpful for predicting perception and response to sounds that are expected to be somewhat masked in the environments in which they operate. This includes the specific example used to examine the metrics in this work, that of UAS (or other AAM vehicles) operating in urban environments.

6. Conclusions

A newly-formulated extension to the SHM loudness has been described, which models the effect of partial masking. The SHM-based partial loudness has been shown to perform similarly to the existing Moore–Glasberg partial loudness model for the cases tested, both in terms of objective predictions and agreement with subjective noticeability data. A key advantage of the SHM-based formulation is the enablement of new P-SQMs to be developed, which offer potential for advanced predictive assessment of perception and response to sound sources operating in real acoustic environments.

Supplementary Material

See the supplementary material for further details on the calculations, tabulated parameter values, and additional figures, comprising a signal processing flow diagram, further partial loudness model comparisons for masked UAS and white noise signals, and example P-SQM calculation results.

Acknowledgments

This research was funded by Horizon Europe and UK Research and Innovation (UKRI) under the UK government's Horizon Europe funding guarantee (Grant No. 10061935). M.J.B.L. is also grateful for the collaboration with Dr-Ing Gil Felix Greco in the development of open-source sound quality analysis tools.

Author Declarations

Conflict of Interest

The authors have no conflicts to disclose.

Ethics Approval

The subjective test data were obtained during an experiment that was granted ethics approval by the Ethics Approval Committee for the University of Salford School of Science, Engineering and Environment (ID: 14321). Participants provided informed consent prior to participation.

Data Availability

The data that support the findings of this study are available from the corresponding author upon reasonable request, subject to RefMap project governance approval. The open-source software implementations of the metrics described have been made publicly available (Lotinga, 2026).

References

[1]It is acknowledged that it would be trivial to also establish an S_{part} metric using the Moore–Glasberg model for N'_{part}; however, this would not be true of the proposed P-SQMs associated with the perception of tonality.

Ecma International (**2025**). *ECMA-418-2:2025, Psychoacoustic Metrics for ITT Equipment — Part 2 (Methods for Describing Human Perception Based on the Sottek Hearing Model)* (Ecma International, Geneva, Switzerland).

Fastl, H., and Zwicker, E. (**2007**). *Psychoacoustics: Facts and Models*, 3rd ed. (Springer-Verlag, Berlin, Germany).

Fidell, S., Pearsons, K. S., and Bennett, R. (**1974**). "Prediction of aural detectability of noise signals," Human Factors **16**(4), 373–383.

Glasberg, B. R., and Moore, B. C. J. (**2005**). "Development and evaluation of a model for predicting the audibility of time-varying sounds in the presence of background sounds," J. Audio Eng. Soc. **53**(10), 906–918, available at https://aes2.org/publications/elibrary-page/?id=13391.

Lee, W., Chun, C., Kim, D., and Lee, S. (**2021**). "Modeling and mapping of combined noise annoyance for aircraft and road traffic based on a partial loudness model," Int. J. Environ. Res. Public Health **18**(16), 8724.

Lotinga, M. J. B. (**2026**). "refmap-psychoacoustics (version 2026) [computer program]," https://github.com/acoustics-code-salford/refmap-psychoacoustics (Last viewed 6-10-2026).

Lotinga, M. J. B., Green, M. C., and Torija, A. J. (**2025a**). "Human perception and response to sound from unmanned aircraft systems within ambient acoustic environments," npj Acoustics **1**(1), 2.

Lotinga, M. J. B., Torjussen, M., and Felix Greco, G. (**2025b**). "Verified implementations of the Sottek psychoacoustic Hearing Model standardised sound quality metrics (ECMA-418-2 loudness, tonality and roughness)," in *Proceedings of the 11th EAA Annual European Conference on Acoustics and Noise Control Engineering: Forum Acusticum/Euronoise*, Malaga, Spain (June 23–26).

Moore, B. C. J., Glasberg, B. R., and Baer, T. (**1997**). "A model for the prediction of thresholds, loudness, and partial loudness," J. Audio Eng. Soc. **45**(4), 224–240, available at https://aes2.org/297publications/elibrary-page/?id=10272.

Rafaelof, M., and Wendling, K. (**2021**). "An algorithm for statistical audibility prediction (SAP) of an arbitrary signal in the presence of noise," J. Audio Eng. Soc. **69**(9), 672–682.

Rizzi, S. A., Christian, A., Letica, S. J., and Lympany, S. V. (**2026**). "Annoyance model assessments of urban air mobility vehicle operations," J. Aircraft **63**(1), 62–77.

Schlittenlacher, J., and Moore, B. C. J. (**2021**). "Temporal integration of partial loudness of helicopter-like sounds," in *Proceedings of the 50th International Congress and Exposition on Noise Control Engineering*, Vol. **263**, pp. 4767–4772.

Sottek, R. (**1993**). "Modelle zur Signalverarbeitung im menschlichen Gehör" ("Models for signal processing in human hearing"), Doctoral thesis, Rhenish-Westphalian Technical University of Aachen, Aachen, Germany.

Sottek, R. (**2016**). "A hearing model approach to time-varying loudness," Acta Acust. united Ac. **102**(4), 725–744.

Sottek, R., Lobato, T., Bender, M., and Becker, J. (**2023**). "Modeling the ISO 226:2023 equal-loudness-level contours by standardized loudness methods," in *Proceedings of Forum Acusticum 2023 10th Convention of the European Acoustics Association*.

Swift, S. H., and Gee, K. L. (**2017**). "Extending sharpness calculation for an alternative loudness metric input," J. Acoust. Soc. Am. **142**(6), EL549–EL554.

Ward, D. (**2017**). "Applications of loudness models in audio engineering," Doctoral thesis, Birmingham City University, Birmingham, UK.

Ward, D. (**2019**). "loudness (version 2019) [computer program]," https://github.com/deeuu/loudness (Last viewed 6-10-2026).