forum acusticum 2026
8-12 September · Graz, Austria

BEYOND NOISE CONTOURS: SIMULATING PSYCHOACOUSTIC METRICS AND POPULATION-BASED NOISE IMPACTS OF INDIVIDUAL FLIGHTS

Stefan Schucker[1], Beat Schäffer[1], Jean Marc Wunderli[1]

[1] *Empa, Swiss Federal Laboratories for Materials Science and Technology, Dübendorf, Switzerland*

Abstract

Noise maps of individual flights conventionally show contours of A-weighted levels: they mark where a given exposure is exceeded, but not how the noise is perceived nor how many people are affected. To move beyond this, the aircraft noise simulation model sonAIR was extended with the psychoacoustic single-event metrics Effective Perceived Noise Level (EPNL) and Zwicker's loudness, and with exposure-response-based population impact indices, all computed from simulated one-third-octave band spectra (25 Hz to 5 kHz, 1 s resolution). EPNL, normally obtained from certification measurements, is computed after realigning the band grid, with a dampened spectral roll-off above 5 kHz that avoids spurious tone-correction penalties; the time-varying loudness is approximated by evaluating the stationary Zwicker method per time step. Verification against calibrated microphone recordings of two A320neo flyovers shows that the simulated metrics are as accurate as the underlying, previously validated emission and propagation model. Linking the resulting maps with population data yields maps of highly annoyed people and expected awakening reactions. Applied to a conventional and a low-noise approach procedure from the SESAR project DYN-MARS, the absolute number of affected people differs strongly between metrics, whereas the relative benefit of the low-noise procedure remains stable.

Keywords: *aircraft noise, psychoacoustic metrics, noise mapping, exposure-response, population impact*

1. Introduction

Aircraft noise remains a major environmental burden: in Europe alone, millions of people are exposed to aircraft noise above the regulatory thresholds [1]. One element of noise mitigation is the development of low-noise flight procedures, which several recent European research initiatives have addressed (e.g. SESAR DYN-CAT, ALBATROSS, DYN-MARS, Stepless). These projects share a common evaluation problem: the potential benefit of a procedure must be quantified by comparing *individual* flights with and without the proposed solution. Standard aircraft noise metrics, however, are based on long-term average sound pressure levels (e.g. L_{den}) and are only applicable to entire operational scenarios.

For single events, the established representation is the noise footprint: a map of contours of the A-weighted sound exposure level L_{AE} or the maximum sound pressure level L_{Amax}, both directly available from simulation. Such contours mark where a given exposure is exceeded, but not how the noise is perceived, nor whether anyone lives there. Two extensions are needed to move beyond them. First, metrics that represent human perception more closely: the Effective Perceived Noise Level (EPNL) [2] and Zwicker's loudness [3]. These are traditionally computed from *measured* signals, EPNL from certification measurements at a few defined positions, loudness from audio recordings. Computing them instead from a *simulation* makes them available at every point of a receiver grid and for procedures that have not been flown, but requires adaptations, because the simulated spectra are band-limited and temporally coarse. Second, exposure-response relationships that translate the mapped exposure into the expected number of affected residents, turning the noise map into an impact map.

This contribution describes both steps for the aircraft noise simulation model sonAIR [4]: the band-grid realignment and high-frequency extrapolation required for EPNL (Section 3), the approximation of the time-varying loudness by a quasi-stationary evaluation (Section 4), the verification of both against calibrated microphone recordings of real flyovers (Sec-

Corresponding author: *stefan.schucker@empa.ch.*

tion 5), and the combination with exposure-response relationships and population data into maps of highly annoyed people and awakening reactions, applied to a conventional and a low-noise approach procedure from the SESAR project DYN-MARS (Section 6).

2. Simulation framework

sonAIR is a spectral, time-step based aircraft noise simulation model [4], [5]. Individual flights are simulated with a temporal resolution of 1 s; the sound emission, including a 3D directivity pattern, accounts for the position and orientation of the aircraft as well as its power setting and configuration (flaps, speed brakes, landing gears). Sound propagation includes terrain and local ground cover. For the present application, the exposure is simulated either on a receiver grid with 100 m mesh size or at individual receiver points (e.g. microphone positions). The output relevant here is, per receiver, a time history of one-third-octave band spectra in 24 bands from 25 Hz to 5 kHz at 1 s steps. Emission and propagation modelling of sonAIR have been validated extensively against measurements [6]; the verification in Section 5 therefore focuses on the newly implemented metrics.

3. EPNL as a simulated quantity

The EPNL implementation follows ICAO Annex 16 [2]: for each time step, the perceived noisiness of the one-third-octave band spectrum is obtained through the noy weighting and aggregated to the Perceived Noise Level (PNL); a tone correction penalises spectral irregularities (up to +6.7 dB), yielding the tone-corrected level PNLT; the maximum PNLTM (including the band-sharing correction) and the duration correction over the 10-dB-down window finally give the EPNL. The chain operates directly on the simulated band spectra, with the 1 s simulation step in place of the 0.5 s step of certification measurements.

Two adaptations are required at the seam between the simulation and the metric. First, the band grids differ: the noy tables are defined for 24 bands from 50 Hz to 10 kHz, whereas sonAIR resolves 24 bands from 25 Hz to 5 kHz. The three bands below 50 Hz are discarded. Second, the three missing bands above 5 kHz (6.3, 8 and 10 kHz) must be filled. Setting them to an inert level (“zero-padding”) creates an artificial spectral edge at 5 kHz that the tone-correction procedure misinterprets as a strong tonal component: at the showcase microphone of Fig. 1, zero-padding raises the EPNL by 2.6 dB through this spurious penalty alone. Instead, the missing bands are extrapolated by continuing the spectral slope of the two highest resolved bands, limited to a minimum roll-off of 3 dB per band (Fig. 1a). With this dampened roll-off, the simulated EPNL reproduces the full-band EPNL computed from the measured signal (Section 5); at the microphone of Fig. 1 the difference is below 0.1 dB.

Figure 1b compares the resulting PNLT time histories of simulation and measurement for an A320neo departure: the simulated history follows the measured one closely over the whole flyover, and in particular within the 10-dB-down window that dominates the duration correction. The remaining differences reflect the emission and propagation modelling rather than the metric implementation.

4. Quasi-stationary loudness

ISO 532-1 [3] defines two procedures: the stationary method (method 1), which operates on a one-third-octave band spectrum, and the time-varying method (method 2), which requires the audio waveform sampled at full rate, and this a band-spectrum simulation cannot provide. The implementation in sonAIR therefore approximates the time-varying loudness quasi-stationarily: because the received signal is effectively stationary within a single 1 s simulation step, the stationary method is evaluated once per time step on the instantaneous receiver spectrum. This yields a loudness time history $N(t)$ in sones and the specific-loudness pattern $N'(z)$ over the Bark scale, from which the maximum loudness N_{max} and the 5-%-exceedance loudness N_5 are derived. The 24 simulated bands map directly onto the lower bands of the 28-band Zwicker grid; the four bands above 5 kHz are set to a negligible level. The numeric core of the stationary method was verified bit-exact against the open-source MOSQITO implementation [7].

Figure 2 characterises the approximation against the time-varying method, both evaluated from the same measured flyover signal. The quasi-stationary loudness tracks the time-varying loudness closely over the entire flyover. The percentile statistic N_5 is robust against the choice of the time step, deviating by less than about 5 % from the time-varying value for time steps of 1 s and below. The maximum N_{max} is systematically underestimated, because averaging over a time step flattens the peak: by about 10 % at the 1 s simulation step, converging to about −3.5 % at a time step of 0.125 s. The remaining offset is caused by the temporal integration (post-masking) of the time-varying method, which the quasi-stationary evaluation cannot represent. For the intended application, the *comparison* of flights simulated with identical settings, this bias cancels; where simulation is compared against measurement (Section 5), both sides are evaluated quasi-stationarily at the same time step.

Both metrics are computed for every grid point of a flyover footprint, so the computational cost matters. The Python reference implementations are fully vectorised over receivers and time steps; for grid-scale runs, both metrics were additionally ported to a parallelised C routine that evaluates them on the fly during the immission calculation, without storing the spectra. On a test grid of 3×10^5 receivers and a 14-minute approach, the complete run including EPNL and N_{max} takes about 90 s with the C routine versus about 20 min with the Python path, a speed-up of roughly a factor of 14. In production, an approach over a 100 m grid of 1.8×10^6 receivers (of which about 4×10^5 lie within reach of the trajectory) is simulated

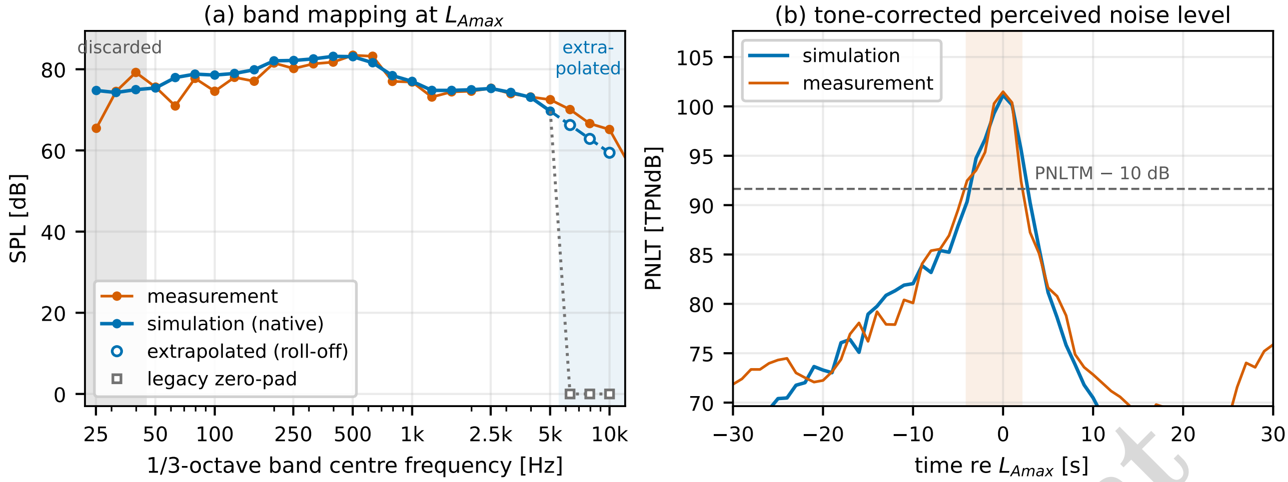

Figure 1: EPNL from simulated spectra, A320neo departure at one microphone: (a) one-third-octave spectrum at the moment of L_{Amax}; the sonAIR bands below 50 Hz are discarded, the bands above 5 kHz are extrapolated with a dampened roll-off (open circles), avoiding the spectral edge of the legacy zero-padding (squares); (b) resulting tone-corrected perceived noise level PNLT over time, simulation versus measurement, with the 10-dB-down duration-correction window of the measurement (shaded).

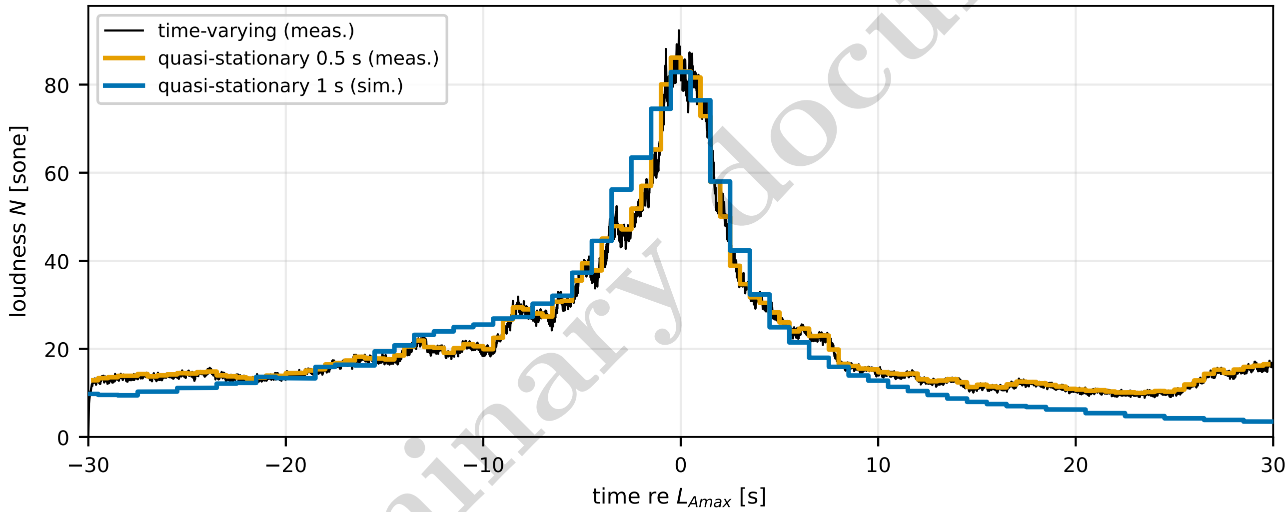

Figure 2: Quasi-stationary approximation of the time-varying loudness, same flyover as Fig. 1: loudness time history $N(t)$ of the time-varying method (ISO 532-1 method 2, computed from the measured waveform) compared with the quasi-stationary evaluation of the measurement (0.5 s spectra) and of the simulation (1 s spectra).

in 30 to 50 s per flight, which makes whole-scenario evaluations with these metrics routinely feasible.

5. Verification against measurements

The implementation was verified against calibrated microphone recordings from a measurement campaign at Zurich Airport (2022). Two flights of an Airbus A320neo (an arrival on runway 34 and a departure on runway 28) were simulated from their flight-data-recorder trajectories, with the exposure evaluated at the eleven microphone positions. The measured signals were processed with the identical EPNL and loudness routines, so that any remaining difference reflects the emission and propagation modelling rather than the metric implementation.

Figure 3 summarises the comparison. With two flights, no statistical statement is intended; the comparison is a verification of the implementation, not a validation of the model. The decisive observation is that, at each microphone, the deviations of EPNL and N_{max} closely follow those of the conventional A-weighted levels: microphones at which L_{Amax} and L_{AE} agree well also agree in the perception-based metrics, and the single microphone in an acoustically shadowed position deviates consistently in all four metrics, indicating a local propagation effect rather than an error of the metric implementation. The slight tendency of the measured loudness to exceed the simulated one is the counterpart of the band limit discussed in Section 3: the measurement contains signal energy above 5 kHz

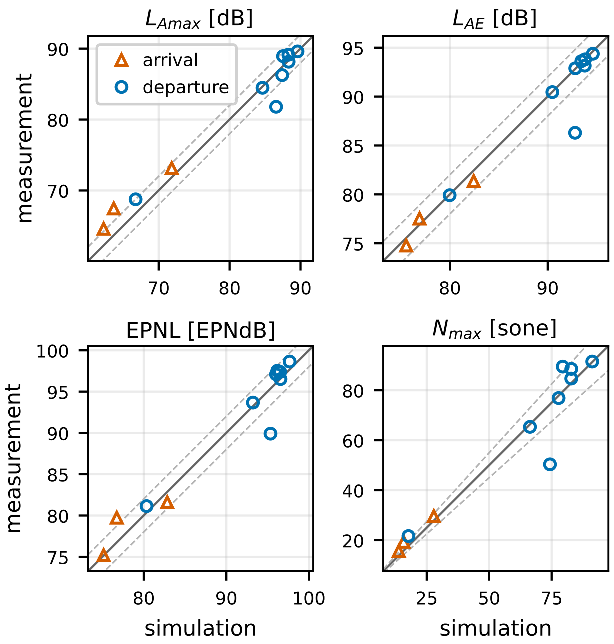

Figure 3: Simulated versus measured single-event metrics at all eleven microphones of the two verification flights (A320neo arrival RWY 34 and departure RWY 28). Dashed guides: ±2 dB (±10 % for N_{max}). The outlier is a microphone in an acoustically shadowed position.

that the simulation does not resolve, which affects the loudness more than the A-weighted levels. Overall, the newly implemented metrics are simulated as accurately as the sound emission and propagation themselves [6]: the metric implementation adds no error of its own.

6. From noise contours to population impact

For the assessment of low-noise procedures, the single-event metrics are combined with exposure-response relationships and a population density map (100 m grid). The probability of a person feeling highly annoyed by a single flyover is modelled as a logistic function of L_{AE}, EPNL or the maximum loudness level, re-analysed from laboratory listening experiments [8]; the probability of an awakening reaction follows the field-study curve of Basner et al. [9] as a function of L_{Amax}, reduced by 15 dB to represent a bedroom behind a tilted window [10]. Multiplying the probabilities with the number of residents per grid cell and summing over the footprint yields the expected number of highly annoyed people and of awakening reactions per flight: an effect-oriented assessment in the spirit of established aircraft noise indices [11], [12], transferred to single events.

Figure 4 makes this chain visible on the map for the reference approach of the application case: the conventional L_{AE} contour footprint (a) is symmetric along the flight path and marks exposure irrespective of land use; the population density (b) is highly inhomogeneous; their combination (c) concentrates the impact over the settlements actually overflown. The noise map thus becomes an impact map that shows *which* communities carry the burden, information that no contour map provides.

Table 1: Population impact of the two approaches of Fig. 5 (relative difference solution vs. reference).

Metric	Ref.	Sol.	Rel.
Highly annoyed (L_{AE})	83 579	79 527	−4.8 %
Highly annoyed (EPNL)	53 620	51 451	−4.0 %
Highly annoyed ($L_{N,max}$)	33 630	32 270	−4.0 %
Awakenings (L_{Amax})	6 457	6 049	−6.3 %

The comparison of two procedures adds a second dimension: Fig. 5 resolves the impact along the ground track for the application case from the SESAR project DYN-MARS, a conventional radar-vectoring arrival (reference, as flown in today's operations) versus a novel FMS-based trajectory optimised for fuel and noise (solution), both flown by the same aircraft type (A320 family) at Zurich Airport. Both population outcomes are generated almost entirely within the final ∼10 NM before the runway, where the aircraft is low over the densely populated area around the airport, and the reduction achieved by the solution is concentrated there as well. This representation links the operational procedure directly to the impact it generates, which is the practical value of the population-based mapping for procedure design.

Table 1 summarises the totals. The absolute number of affected people differs by nearly a factor of three between the annoyance metrics, reflecting the different exposure-response curves; the relative benefit of the solution, however, remains stable within −4.0 to −6.3 %. The choice of metric thus strongly affects the headline count, but not the conclusion which procedure is preferable. For the investigated conventional jet aircraft, EPNL and loudness correlate strongly with L_{AE} across the footprint (broadband-dominated emission); their added value is expected to grow for sources with tonal components such as propeller aircraft, helicopters or drones.

7. Conclusions

The aircraft noise simulation model sonAIR was extended with the psychoacoustic single-event metrics EPNL and Zwicker loudness, computed from simulated one-third-octave band time histories instead of measured signals. Two implementation choices proved essential: (i) for EPNL, a dampened spectral roll-off above the 5 kHz band limit of the simulation, which avoids spurious tone corrections of up to 2.6 dB and reproduces the full-band measured EPNL; and (ii) for loudness, a quasi-stationary evaluation of the stationary Zwicker method per simulation time step, whose percentile statistic N_5 is robust to the time step, while the peak statistic N_{max} carries a known, systematic

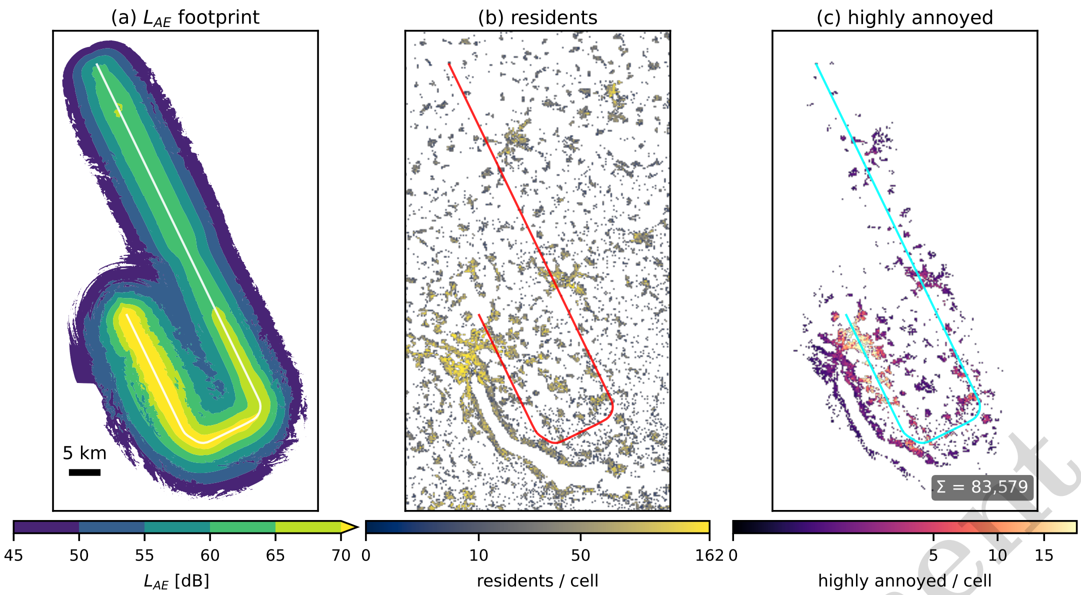

Figure 4: From the noise contour to the population impact, for a conventional radar-vectoring approach to Zurich Airport (ground track in the panels): (a) L_{AE} noise-contour footprint; (b) residents per 100 m grid cell; (c) resulting map of highly annoyed people per cell, the product of the exposure-response probability from (a) and the population from (b).

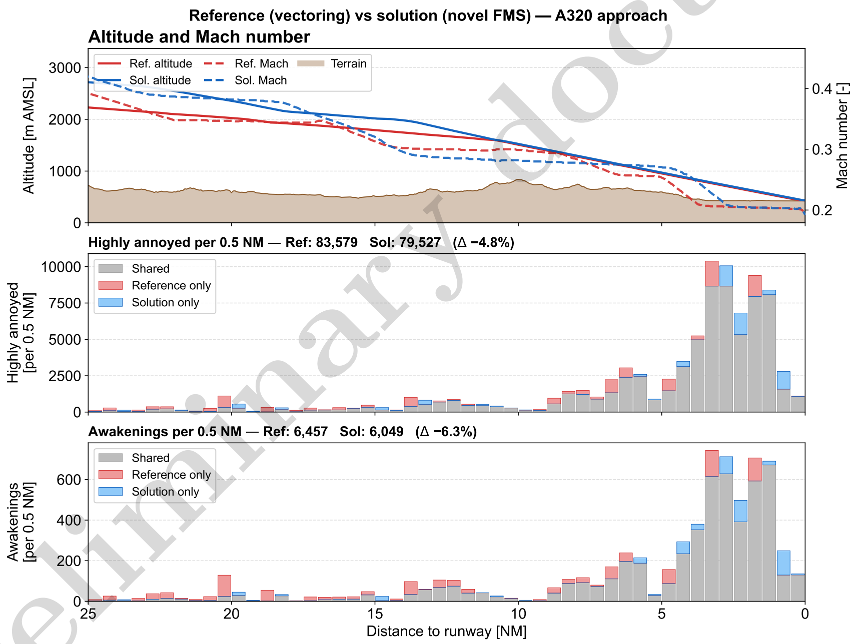

Figure 5: Population impact resolved along the flight path: reference (conventional radar vectoring, red) versus solution (novel FMS-based procedure, blue) as a function of the distance to the runway. Top: altitude profile with terrain (solid) and Mach number (dashed). Middle and bottom: highly annoyed people (from L_{AE}) and awakening reactions (from L_{Amax}) per 0.5 NM slice, split into the contribution shared by both flights and the parts unique to each.

and therefore comparable underestimation. Verification against calibrated recordings of two A320neo flyovers confirms both implementations within the accuracy of the underlying emission and propagation model. Combining the simulated metrics with exposure-response relationships and population data moves the assessment beyond noise contours: the resulting maps of highly annoyed people and awakening reactions localise the impact over the communities actually affected, resolve it along the flight path, and enable an effect-oriented comparison of individual flights, in which the relative benefit of a low-noise procedure proves stable across metrics, while the absolute counts do not.

8. Acknowledgments

This research was funded by the State Secretariat for Education, Research and Innovation SERI (grant no. 23.00355) as part of the SESAR project DYN-MARS (DYNamic Management of Aircraft Configuration and Route Structures).

References

[1] European Environment Agency, "Environmental noise in Europe 2025," European Environment Agency, Tech. Rep., 2025. DOI: 10.2800/9200716

[2] ICAO, *Annex 16 to the convention on international civil aviation: Environmental protection, volume I – aircraft noise*, 8th ed., International Civil Aviation Organization, 2017.

[3] International Organization for Standardization, *ISO 532-1:2017 acoustics – methods for calculating loudness – part 1: Zwicker method*, International Organization for Standardization, Geneva, 2017.

[4] J. M. Wunderli et al., "sonAIR – a GIS-integrated spectral aircraft noise simulation tool for single flight prediction and noise mapping," *Acta Acustica united with Acustica*, vol. 104, pp. 440–451, 2018. DOI: 10.3813/AAA.919180

[5] C. Zellmann, B. Schäffer, J. M. Wunderli, U. Isermann, and C. O. Paschereit, "Aircraft noise emission model accounting for aircraft flight parameters," *Journal of Aircraft*, vol. 55, no. 2, pp. 682–695, 2018. DOI: 10.2514/1.C034275

[6] D. Jäger, C. Zellmann, F. Schlatter, and J. M. Wunderli, "Validation of the sonAIR aircraft noise simulation model," *Noise Mapping*, vol. 8, no. 1, pp. 95–107, 2021. DOI: 10.1515/noise-2021-0007

[7] Green Forge Coop, *MOSQITO: A unified and modular sound quality metrics library*, Version 1.2.1, 2024. DOI: 10.5281/zenodo.5284054

[8] F. Georgiou, B. Schäffer, A. Heusser, and R. Pieren, "Predicting aircraft noise annoyance from single flyby events using acoustical and psychoacoustic metrics," *The Journal of the Acoustical Society of America*, vol. 159, no. 2, pp. 1707–1720, 2026. DOI: 10.1121/10.0042533

[9] M. Basner, A. Samel, and U. Isermann, "Aircraft noise effects on sleep: Application of the results of a large polysomnographic field study," *The Journal of the Acoustical Society of America*, vol. 119, no. 5, pp. 2772–2784, 2006. DOI: 10.1121/1.2184247

[10] B. Locher et al., "Differences between outdoor and indoor sound levels for open, tilted, and closed windows," *International Journal of Environmental Research and Public Health*, vol. 15, no. 1, p. 149, 2018. DOI: 10.3390/ijerph15010149

[11] M. Brink, D. Schreckenberg, G. Thomann, and M. Basner, "Aircraft noise indexes for effect oriented noise assessment," *Acta Acustica united with Acustica*, vol. 96, no. 6, pp. 1012–1025, 2010. DOI: 10.3813/aaa.918364

[12] B. Schäffer, G. Thomann, P. Huber, M. Brink, S. Plüss, and R. Hofmann, "Zurich aircraft noise index: An index for the assessment and analysis of the effects of aircraft noise on the population," *Acta Acustica united with Acustica*, vol. 98, no. 3, pp. 505–519, 2012. DOI: 10.3813/aaa.918533