

RESEARCH ARTICLE

SANE: An index of anthropogenic noise levels for wildlife research in terrestrial ecosystems

Matteo Giuliani  | Davide Mirante  | Luca Francesco Russo  | Andrea Zampetti  | Luca Santini 

Department of Biology and
Biotechnologies 'Charles Darwin',
Sapienza University of Rome, Rome, Italy

Correspondence

Matteo Giuliani

Email: matteo.giuliani@uniroma1.it;
giuliani.matte@gmail.com

Funding information

Italian Ministry for Universities and
Research (MUR), Grant/Award Number:
B53D23012400001

Handling Editor: Alison Johnston

Abstract

1. Understanding the impact of noise pollution on wildlife is challenging due to the complexity of isolating the anthropogenic components from the soundscape. Soundscape studies usually employ acoustic indices that apply thresholds to isolate anthrophony from natural sounds. However, natural and anthropogenic sounds do not always conform to these thresholds, hindering both accuracy and comparability of the resulting indices. While significant progress has been made in the automated identification of acoustic events through artificial intelligence-based classifiers, an effective method to quantify anthropogenic acoustic pressure is still lacking.
2. We propose a novel index, the Selective Anthropogenic Noise Exposure (SANE), which leverages BirdNET deep neural network to isolate human-related sounds from recordings. The index consists of the sum of the Median Amplitude Index of all human noise categories, thereby capturing both the intensity and cumulative impact of multiple disturbance events, while also enabling the decomposition of noise level across different categories of disturbance (e.g. traffic, human voices).
3. We test SANE's performance in a real urban setting and through soundscape simulations and compare it with two frequency-based indices and another artificial intelligence-based acoustic index (CityAnthroNet). SANE was effective in describing human noises within areas characterized by different levels of urbanization, improving upon other indices' shortcomings. Additionally, SANE was robust at very fine temporal scales, precisely quantifying anthrophony levels for single recordings. Among the indices considered, SANE was the only one that remained insensitive to low-frequency biophony, which confounded both frequency and artificial intelligence-based metrics.
4. By leveraging artificial intelligence-based classifiers to detect multiple human-made sound classes, the SANE index captures the intensity of anthropogenic noise. At the same time, natural sounds falling within the traditionally identified anthrophony range are ignored. Furthermore, SANE has the potential to assess the relative contribution of different noise types to overall acoustic pollution,

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2026 The Author(s). *Methods in Ecology and Evolution* published by John Wiley & Sons Ltd on behalf of British Ecological Society.

opening new research avenues regarding the effects of acoustic pollution on wildlife.

KEYWORDS

acoustic indices, acoustic monitoring, anthropogenic disturbance, impact assessment, noise pollution, soundscape analysis

1 | INTRODUCTION

An increasing number of studies have highlighted negative effects of anthropogenic noise (anthrophony) on wild fauna, ranging from behavioural to physiological alteration (Brown et al., 2012; Fuller et al., 2007; Shannon et al., 2015; Slabbekoorn & Peet, 2003). Effects include the disruption of predator–prey interactions (Kaiser & Hammers, 2009; Sweet et al., 2022), changes in communication efficiency (Lampe et al., 2014; Loss et al., 2021) and ultimately consequences for individual fitness (Cruz et al., 2018). With the projected increase in human population and related infrastructures, anthrophony levels are expected to increase in the near future (Hildebrand, 2009), underscoring the need for robust tools to quantify and monitor such disturbances.

Thanks to the recent developments of acoustic sensors, passive acoustic monitoring has recently gained popularity in ecology (Alcocer et al., 2022). The subsequent urge to efficiently quantify acoustic phenomena has led to the emergence of several acoustic indices, which are mathematical tools that summarize, into simple numbers, specific soundscape features. Despite acoustic indices proved to be useful for characterizing some soundscape components and their temporal patterns (Bradfer-Lawrence et al., 2019), their use for quantifying biodiversity has faced sharp criticism (Giuliani et al., 2024; Sugai et al., 2026). Among the numerous existing indices, some have been specifically designed to quantify anthrophony, that is, the Power Spectral Density (PSD) calculated between 1 and 2 kHz, and to measure the interaction between human noises and biophony, that is, the Normalized Spectral Dissimilarity index (NDSI; Kasten et al., 2012). These indices rely on theoretical frequency boundaries that delimitate biophony (2–11 kHz) and anthrophony (1–2 kHz) (Pijanowski et al., 2011). However, soundscape components are not always confined to these boundaries, with geophony (i.e. abiotic sounds) lacking standardized frequency boundaries (Putland et al., 2017), and many human and biological sounds exceeding the stated limits (Jacewicz et al., 2023; Sánchez-Giraldo et al., 2020).

Anthropogenic sounds can be highly diverse and produce different effects on species depending on their frequency, intensity and temporal pattern (Engel et al., 2024; Gomes et al., 2022). For example, anthropogenic noise that spectrally overlaps with biologically relevant signals masks communication or predator cues, forcing animals to adjust their vocalizations' frequency or increasing perceived predation risk (Engel et al., 2024; Higham et al., 2021). Acute, high-amplitude sounds can evoke immediate responses in some

animals (e.g. anti-predatory responses; Randler, 2006), whereas low-intensity chronic noise often leads to prolonged changes in vigilance or activity budgets without necessarily triggering escape behaviours (Collins et al., 2022). On the other hand, biophony encompasses a huge variety of sounds, some of which fall within the anthrophony range, for example, those produced by the Eurasian Wild Boar or by the Yellow-legged Gull. As a result, there is a spectral overlap between biological and anthropogenic sounds, which causes an ambiguity in distinguishing the sounds' nature based solely on frequency range. The direct result is the limitation in the effectiveness of acoustic indices based on fixed frequency thresholds. In this context, advanced methods for automated sound analysis can be a valuable support. Artificial Intelligence (AI) is increasingly applied in ecology to support and accelerate data processing (Pollock et al., 2025). In the field of bioacoustics, several algorithms based on Neural Networks have been applied to identify species from acoustic recordings (e.g. BirdNET; Kahl et al., 2021) or to discriminate between anthrophony and biophony (e.g. CityNet; Fairbrass et al., 2019). In particular, CityNet provides an estimate of the proportion of anthropogenic and biological sounds across single recordings by leveraging two different classifiers (Fairbrass et al., 2019), one specific for human noise (CityAnthroNet) and one specific for biological sounds (CityBioNet). However, being designed to quantify the proportion of time occupied by anthrophony, CityAnthroNet does not account for acoustic intensity; this latter has been proven to be an important driver of behavioural disturbance and physiological stress (Shannon et al., 2015).

Here, we present a novel index, the SANE, which combines AI to isolate anthropogenic noise from acoustic data and a simple acoustic index to estimate anthropogenic noise levels. SANE results in a detailed measurement of anthrophony, capturing the nature and the cumulative intensity of anthropogenic noises in the soundscape. Moreover, SANE proved robust to non-target sounds, which can determine inconsistencies in the performance and usefulness of traditional acoustic indices (Sugai et al., 2026). We aim to (i) illustrate the workflow to produce the SANE index, (ii) assess the ability of SANE to capture spatial and temporal patterns of anthropogenic noise along an anthropogenic gradient, (iii) evaluate its decomposability into distinct human noise categories and (iv) compare its performance with other existing indices.

Specifically, we illustrate how SANE performs in describing the intensity and identity of anthropogenic noise in two case studies and compare it with the anthrophony level estimated by CityAnthroNet (Fairbrass et al., 2019), and with frequency-based indices, that is,

NDSI and PSD (1–2 kHz). First, we used an extensive passive acoustic monitoring conducted in the city of Rome (Italy), spanning from large peripheral parks to the urban centre, to visualize the spatial and temporal pattern of all calculated indices. Then, we evaluated their consistency and overall performance in capturing anthrophony levels. Furthermore, we showcase SANE's decomposability by using the values obtained for each human noise category, highlighting their relative temporal and spatial variation along the urbanization gradient.

Finally, we test indices' abilities to describe anthrophony using simulated soundscapes mimicking gradients of noise pollution and varying proportions of bird species vocalizing within the range of biophony and anthrophony, with the aim of evaluating their robustness under controlled acoustic conditions.

2 | METHODS

2.1 | BirdNET as a classifier of human noise

BirdNET is a Deep Neural Network originally designed for automated birds' species identification, which is widely used in ecological studies using acoustic data (e.g. Paterson et al., 2024; Plat et al., 2025; etc.). In addition to ~6000 bird species, BirdNET V2.4 is also capable of recognizing nine classes of anthropogenic sounds: specifically, human vocal, human non-vocal (i.e. sounds directly produced by humans, like sneezes or coughing), human whistle, engine, power tools, fireworks, siren, guns and dogs. The human noise training dataset consisted of almost 8000 recordings derived from the Google AudioSet (Gemmeke et al., 2017), the Freefield1010 (Stowell & Plumbley, 2013) and the WarbLR datasets, which contain recordings from around the world (Kahl et al., 2021).

Other deep learning systems have been specifically designed to quantify anthropogenic noise, like CityAnthroNet (Fairbrass et al., 2019). However, it was trained exclusively on recordings from the Greater London area (United Kingdom) and does not distinguish disturbance classes, returning only the proportion of human noise activity across the recording without estimating its intensity.

2.2 | SANE index calculation

The SANE is calculated in three steps (Figure 1). Firstly, acoustic files are processed by BirdNET to identify human-related sounds. BirdNET divides the audio file into snippets of 3 s and classifies each in a sound category. So, all snippets classified as human-related sounds are identified. BirdNET also provides a confidence score associated with each sound. Based on a local validation (see Section 2.3.2), an optimal confidence threshold should then be selected to achieve the desired precision. For each 3-second audio snippet classified as anthropogenic by BirdNET and exceeding the selected confidence threshold, the corresponding audio segment is extracted from the original waveform using the start and end times provided by the classifier. Each extracted segment is then band-pass filtered over the audible frequency range (20–20,000 Hz) using a fourth-order Butterworth filter to remove very low- and very high-frequency components.

Then, the Median Amplitude index (Depraetere et al., 2012) is subsequently computed in R using the *M()* function from the *seewave* package (Sueur, Aubin, & Simonis, 2008; Sueur, Farina, et al., 2008). The index corresponds to the median of the amplitude envelope, which is not computationally demanding and quantifies the amount of acoustic energy associated with each anthropogenic acoustic signal identified without being overly sensitive to outliers.

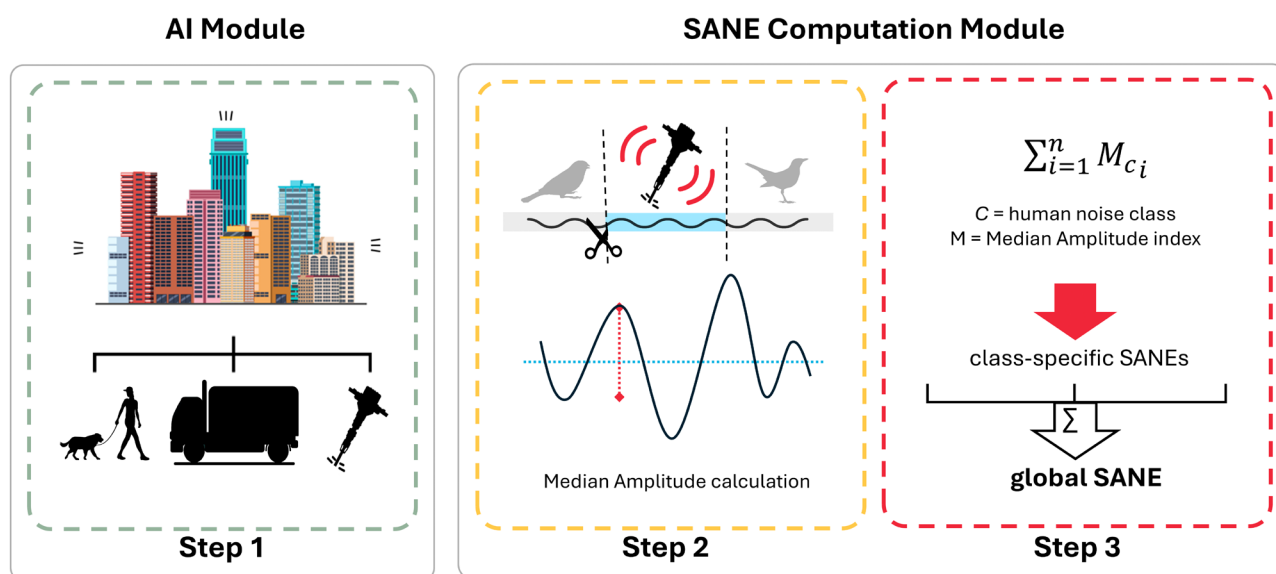


FIGURE 1 Workflow of the SANE index calculation. The first step uses BirdNET to identify target human noise signals. Thus, index processing moves in R: Specifically, in step 2, the BirdNET output is used to compute median amplitude for each signal; in step 3, median amplitude values are summed within and across noise classes to obtain the SANE indices of interest.

Lastly, Median Amplitude Index values estimated from each snippet are summed for each audio to produce a SANE value for each human noise category, as well as a global SANE value considering all human noise categories. If no human noise was recorded throughout the entire recording, the SANE index equals zero. The described workflow is embedded in an R customizable function, which is provided through a GitHub repository (<https://github.com/matpagle/sane>). To assess the robustness of the SANE index relative to alternative approaches and illustrate its application, we apply it both to real-world recordings (2.3) and to controlled simulated soundscapes (2.4).

2.3 | Real soundscape tests

2.3.1 | Acoustic survey

To illustrate the application of the SANE index, we leveraged an extensive passive acoustic monitoring campaign conducted from March to June 2024 in the city of Rome (Italy). The study area is characterized by more than 50 bird species breeding within the city boundaries (Cignini & Zapparoli, 1996). The survey consisted of a total of 75 listening points within the motorway (Grande Raccordo Anulare) encircling the most populated and urbanized part of the city (Figure S1). The minimum distance between each listening point was no less than 100 m to avoid spatial pseudoreplication (Alcocer et al., 2022). To characterize the urban environment, we used the map provided by Russo et al. (2025) that classifies the landscape into three urbanization categories: low urbanization, representing homogenous green space; moderate urbanization, where impervious surfaces are interspersed with small green patches; and high urbanization, representing highly impervious areas.

Each listening point consisted of one AudioMoth, which remained in place one night per month from 5:00 PM to 9:00 AM, recording for 5 minutes every 20 minutes (5 hours and 20 minutes of recording per day). The deployment period was designed to intercept the vocal activity peaks of diurnal and nocturnal birds while minimizing the risk of device theft. We sampled only during favourable meteorological conditions to minimize strong geophony (e.g. wind gusts, heavy rain, strong running water) as a confounding factor in acoustic indices. The final set consisted of a total of 13,100 5-minute recordings, 64 per day of deployment for each listening point.

2.3.2 | BirdNET validation

We used BirdNET to identify the human noise signals in each audio file of all recordings collected (~65,000 minutes). We compiled an ad hoc list comprising the full set of human-related sounds encompassed by BirdNET for audio analysis. BirdNET processes audio files in 3-second snippets, assigning a confidence score when a target sound is identified. We set the minimum confidence score

threshold at 0.1, applied a 2 s of overlap on prediction segments to enhance recall (Pérez-Granados et al., 2025) and maintained the default sensitivity (1.0). To ensure that only target sounds were classified and to avoid errors from false positives, which would affect the index's accuracy, we validated the outputs. We followed the BirdNET validation procedure proposed for birds' sounds by Pérez-Granados et al. (2025). We randomly selected 100 classified 3-second snippets from the entire BirdNET output for each class of anthropogenic noise, ensuring an even distribution across all confidence score intervals (e.g. 0–0.2, 0.2–0.4, ..., 0.8–1), to represent the entire spectrum of confidence. Each segment was manually reviewed and annotated to confirm the presence of anthropogenic noise. Annotations indicated whether the segment truly contained anthropogenic sounds (true positive) or not (false positive), as well as whether the detected sound matched the class assigned. We then computed the precision score, that is, the ratio between true positives and the sum of true and false positives. This was done for each threshold considered (i.e. thresholds ranging from 0 to 1 in steps of 0.05), both for each anthropogenic noise class separately and anthrophony overall. Furthermore, we also assessed the recall, that is, the ability of the classifier to identify true positives within the dataset. Specifically, we randomly selected five recordings that contained at least one recognized signal of a specific class, for a total of 45 5-minutes recordings. Every recording was listened to, and human sounds were manually annotated for each 3-second snippet. Then, we calculated the recall as the ratio between true positives and the sum of true positives and false negatives. Lastly, for completeness, we also computed the F1 scores as a balanced measure of precision and recall, defined as twice the ratio between their product and their sum.

2.3.3 | Anthrophony estimates through SANE and other indices

To calculate SANE indices on our acoustic dataset, we followed the three-step procedure described above (Figure 1). We used the precision metrics of the classifiers to identify the optimal thresholds to avoid false positives. We used precision instead of F1 score for threshold selection to minimize the chance of including non-target sounds in our index. We then obtained one SANE value per sound class and one global SANE value per recording.

For each recording, we also calculated the CityAnthroNet level of anthropogenic acoustic activity (Fairbrass et al., 2019), the Normalized Difference Soundscape Index (NDSI) and the Power Spectral Density (1–2 kHz, PSD) as described by Kasten et al. (2012). CityAnthroNet classifier was run in Python v.2.7.12 (Python Software Foundation, 2016) using scikit-learn v.0.18.1 (Pedregosa et al., 2011) and matplotlib v.1.5.1 (Hunter, 2007), as designed by Fairbrass et al. (2019). The NDSI was defined as the ratio $\frac{[\text{biophony} - \text{anthrophony}]}{[\text{biophony} + \text{anthrophony}]}$ of the estimated PSDs, computed using Welch's method (Welch, 1967), of the anthrophony (1–2 kHz) and biophony (2–11 kHz). Acoustic indices computation

procedures were conducted in R Studio (R Core Team, 2024) using the *seewave* package (Sueur, Aubin, & Simonis, 2008; Sueur, Farina, et al., 2008).

2.4 | Simulated soundscapes tests

To evaluate the ability of acoustic indices in capturing anthropophony while controlling for confounding variables, we produced 100 simple simulated soundscapes consisting of four scenarios, containing only human noises and birds' vocalization. We designed the simulated soundscapes to be temporally structured as a sequence of contiguous, 3-second audio segments with various degrees of temporal overlap and no silent gaps. We retrieved the sound snippets from a small classified acoustic dataset, consisting of ornithological recordings, collected from xeno-canto (<https://xeno-canto.org>), and anthropogenic sounds with low background noise, extracted from the field case study (Section 2.3; see Table S1 for details). For the ornithological snippets, species were randomly selected among Italian birds to represent different habitat types (e.g. agricultural, forest and wetland environments). Only high-quality recordings with minimal background noise and no overlapping vocalizations from other species were retained, and all files were trimmed using Audacity 3.2.2 into non-overlapping 3-second snippets to match the temporal resolution of the classifier. We classified audio segments in our dataset into three classes: avian vocalizations that do not fall into the anthropophony frequency range (hereafter, *non-anthropophonic birds*), birds' signals that overlap with the anthropophony frequency range (hereafter, *anthrophonic birds*) and human noise.

Each of the 100 simulated soundscapes had a total length of 1200seconds, divided into four 300-second phases, each with a distinct snippet class selection: (i) only *non-anthropophonic birds* and *human noise*; (ii) only *anthrophonic birds* and *human noise*; (iii) both *anthrophonic* and *non-anthropophonic birds*, and *human noise*; (iv) only *anthrophonic* and *non-anthropophonic birds*. In the first three phases (i–iii), the proportion of bins filled with *human noise* increases by 1/10 every 30seconds, starting from zero. In the last phase (iv), anthropophonic bird snippets increased progressively, following the previous human noise pattern.

Finally, to generate soundscapes that were not influenced by the source recordings' audio signal intensity, we normalized each snippet to a fixed amplitude (1.0). We produced the simulated soundscapes in R (R Core Team, 2024), using the *tuneR* (Ligges et al., 2023) and the *seewave* (Sueur, Aubin, & Simonis, 2008; Sueur, Farina, et al., 2008) libraries.

As for the real soundscapes, we computed the global SANE indices, using the precision metrics to guide the choice of confidence threshold to achieve the maximum precision possible (1), the CityAnthroNet level of anthropogenic acoustic activity, the NDSI and the PSD (1–2kHz). Unlike the real soundscape recordings, indices for simulated soundscapes were not computed over the entire recording. Instead, they were calculated for each 30-second

segment to reflect the sequential structure of the simulated soundscapes and to enable a fine-scale assessment of index responses to increasing snippet frequencies of human noise (phases i–iii) and anthropophonic bird vocalizations (phase iv). Finally, the mean value of each index was computed for each 30-second segment across the 100 simulated soundscape replicas, and the results were compared. To facilitate visual comparison of index responses across the simulated soundscape gradients, index values were normalized between 0 and 1 prior to plotting.

2.5 | Indices comparison

For real soundscape recordings, we computed eight indices: five SANE indices (a global one and one for each human noise category retained: Dog, Engine, Human Vocal and Siren), CityAnthroNet activity level, PSD (1–2kHz) and NDSI. As a result, we obtained a temporal series of all indices for each site. For temporal series estimation, we grouped sites based on their Urbanization category (see Section 2.1). For direct comparability among indices, we scaled them to a mean of zero and a standard deviation of 1. Furthermore, because higher values of the NDSI indicate lower anthropogenic presence, for enhancing interpretability, we used the negative form of the NDSI (invNDSI). Lastly, we computed Pearson correlation coefficients among acoustic indices for each urbanization category.

For simulated soundscapes, we aimed to evaluate the consistency of each acoustic index in capturing anthropogenic noise under controlled conditions, and we measured their correlation (Pearson coefficients). We summarized the behaviour of each index across simulated soundscapes by calculating the mean value for each 30-second time step.

Our dual approach, combining real and simulated soundscapes, enabled a comprehensive evaluation of index performance. Specifically, it allowed us to assess each index's ability to capture the temporal dynamics of anthropogenic noise in complex environments, to discriminate among urbanization gradients and to maintain consistency across repeated, controlled soundscapes.

3 | RESULTS

3.1 | BirdNET classifier's performance

Within the complete audio material, BirdNET identified a total of 434'865 segments containing human noise. For BirdNET, the most represented human noise class was 'Engine', which amounted to 73% of the total. Conversely, the less frequent class was 'Human Whistle', consisting only of 28 detections. A relatively frequent disturbance class, 'Gun', consisted only of false positives (although still related to other anthropogenic sounds).

BirdNET's output validation process highlighted a notable ability to discriminate between natural (e.g. birds' songs) and anthropogenic

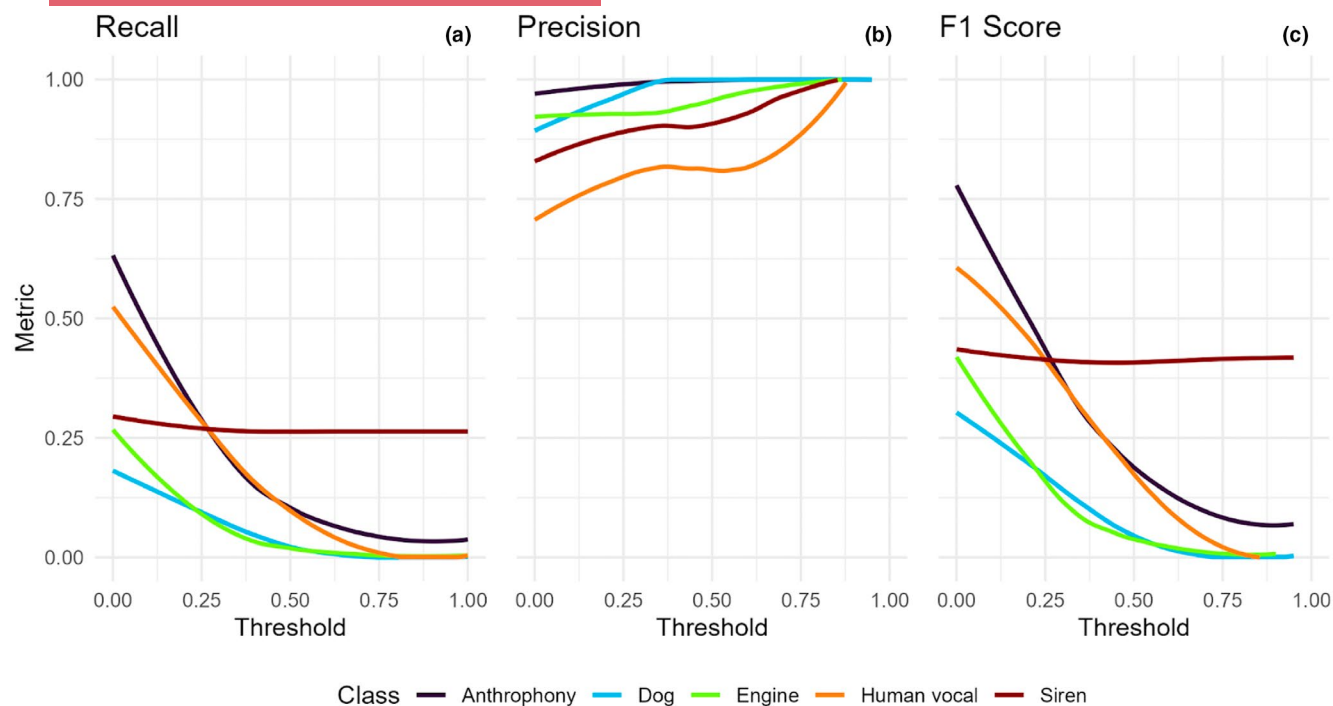


FIGURE 2 Ranges of recall (a), precision (b) and F1-score (c) for selected human noise classes identified by BirdNET across threshold values. Metrics were computed for the four most represented classes (i.e. Dog, Engine, Human Vocal and Siren) as well as for the global Anthropophony category.

noise: the mean precision value throughout the confidence thresholds was 0.99 for BirdNET, with a minimum value of 0.97 for the 0.1 threshold. The minimum confidence threshold that maximized precision for anthropophony was 0.5 (Figure 2).

The performance of BirdNET in classifying human noise as anthropophony varied across classes (Figure 2). Several classes resulted in being quite rare, making the computation of their metrics challenging; thus, we considered them for the global anthropophony but did not compute their precision, recall and F1-score. Classes that were sufficiently represented were Human Vocal, Dog Barking, Engine and Siren.

Recall values varied for all disturbance classes, with the general anthropophony exceeding 0.6 at the 0.1 confidence threshold (Figure 2a). Due to the high precision, the classifier obtained a relatively high F1-score for the general anthropophony at the minimum Confidence threshold possible (Figure 2c).

3.2 | Real soundscape temporal patterns of anthropophony

Frequency-based indices showed high correlation (Table S2) and similar temporal patterns to each other across different levels of urbanization, with PSD displaying larger fluctuation. On the contrary, AI-based indices were markedly distinct, showing low correlations both among themselves and with frequency-based indices. A notable difference is that, as

expected, AI-based indices showed an increase in mean index value with increasing urbanization levels, whereas frequency-based indices did not (Low urbanization=green areas, moderate urbanization=impervious areas interspersed with green areas and high urbanization=impervious areas; Figure 3).

The temporal trends in the four indices also revealed interesting patterns in relation to the increase in human activity and biophony around sunset and sunrise hours, across the three urbanization levels (Figure 3). The SANE was the only index highlighting an increase in intensity in highly urbanized areas during evening hours, a pattern consistent with increased human activity typically occurring in the late afternoon and evening. In contrast, CityAnthroNet showed higher activity in low and moderate urbanization areas than in highly urbanized areas during late hours, a pattern that may be associated with increased biological activity around sunset. Frequency-based indices did not show consistent increases across the three urbanization levels in the early morning. On the contrary, both AI-based indices tended to increase in the morning, for the SANE, which does not in low urbanized areas. Interestingly, the CityAnthroNet activity level increases substantially in the morning, also in the low urbanized areas.

When the SANE index was decomposed into different noise classes (Figure 4), it revealed that low-urbanized areas were dominated by human voices, possibly pointing to recreational activities; low and moderately urbanized areas by traffic noise and dog barking; and moderately to highly urbanized areas by sirens, possibly related to traffic over the main roads.

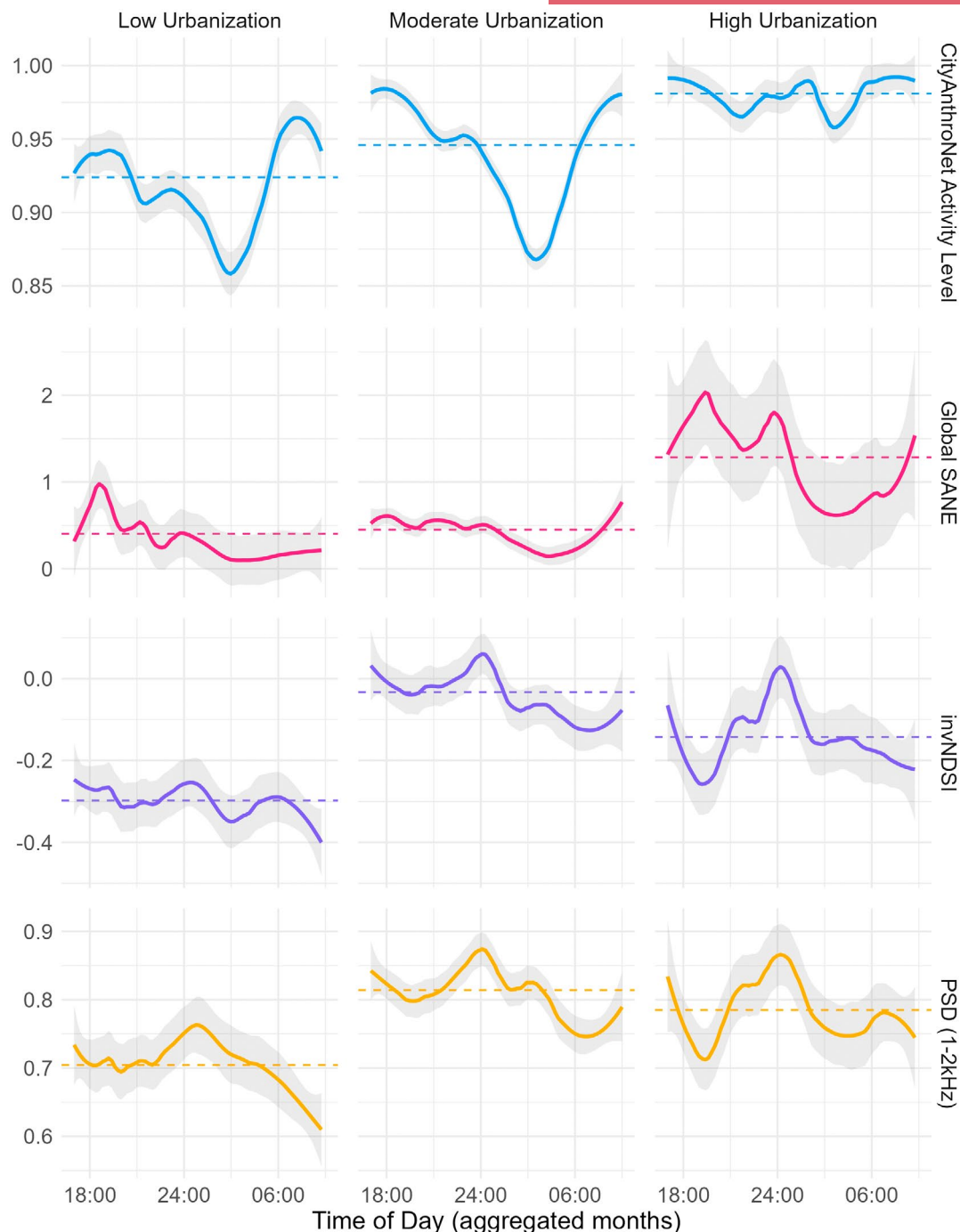


FIGURE 3 Time-series plot of the CityAnthroNet Activity Level, the Global SANE, the invNDSI and the PSD (1–2 kHz), through urbanization cluster, that is, low urbanization (a), moderate urbanization (b) and high urbanization (c), obtained by aggregating together the data from all the sampled months (one recordings per month, 3 months per site) and using the *Loess Regression* approach ($span=0.6$). The shading around the curves represents the 95% confidence intervals. Dotted horizontal lines represent the mean values of each index in each urbanization class. During the sampling period, mean sunrise and sunset times were 05:45 AM and 08:30 PM, respectively. During the sampling period, the mean time for sunrise was 05:45 AM and for sunset 08:30 PM.

3.3 | Simulated soundscape acoustic indices patterns and consistency

The SANE index was the least affected by the presence of anthroponic bird vocalization across the 100 simulated soundscapes. During the first three scenarios (Figure 5i–iii), SANE

values increased proportionally to the number of bins occupied by human noises in each 30-second step; in the last scenario, which only included bird sounds (Figure 5iv), its pattern remained low and flat. In contrast, frequency-based indices were generally higher in the second scenario, including anthroponic bird sounds (Figure 5ii) than in scenarios one and three (Figure 5i, iii), resulting in a less

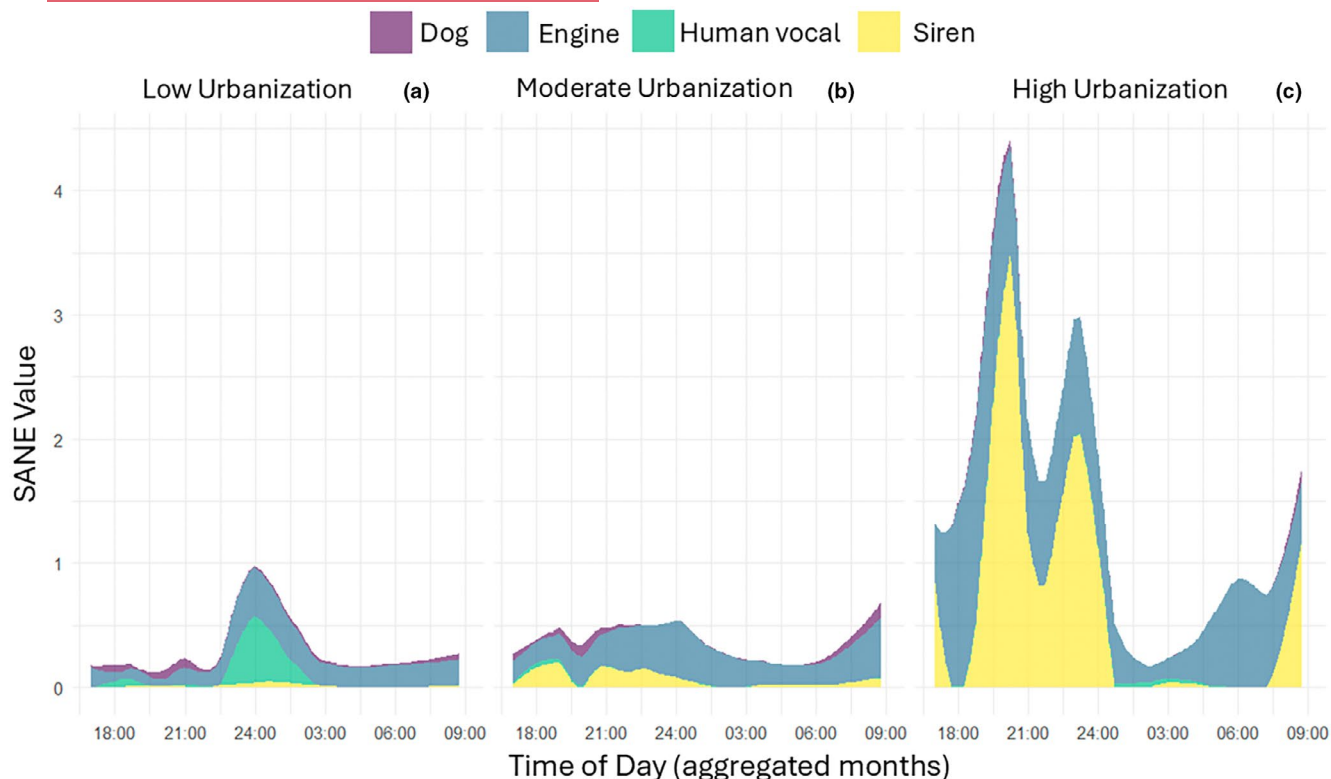


FIGURE 4 Area chart showing the contribution of each human noise class to the global SANE over the time series, aggregated across all sampled months (one recording/month, 3 months/site) for low (a), moderate (b), and high (c) urbanization levels. Temporal trends were smoothed using a Loess approach (span=0.6). Three classes are excluded: 'Gun' (no true positives), 'Human-whistle' (insufficient detections) and 'Human non-vocal' (unable to reach 90% precision). The upper boundary of each area represents the global SANE value. During the sampling period, mean sunrise and sunset times were 05:45 AM and 08:30 PM, respectively.

marked increase relative to the actual rise in anthropogenic noise snippets frequency. In the fourth scenario, both indices increased in response to the growing presence of anthrophonic birds. All indices were highly positively correlated across the first three temporal scenarios (Table S3). However, in the last scenario, SANE exhibited zero variance, consistently taking a value of zero due to the absence of target sounds, preventing the computation of correlation coefficients involving this index.

Surprisingly, the CityAnthroNet anthrophony activity level showed a pattern similar to the frequency-based acoustic indices, being higher in the second scenario, including anthrophonic birds (Figure 5ii) and increasing progressively in the fourth scenario lacking anthropogenic noise (Figure 5iv). In particular, the increase in the last phase was sharper than that of the invNDSI and the PSD (1–2 kHz). Finally, AI-based indices displayed high consistency across the 100 simulated soundscapes (Figure S2a,d), whereas frequency-based indices exhibited high variability under all scenarios (Figure S2b,c).

4 | DISCUSSION

Here, we introduced the SANE, which combines an AI-based classifier with an acoustic index to quantify the cumulative intensity

of anthropogenic noise in terrestrial ecosystems. Deep learning tools make it possible to overcome one of the most important acoustic indices limits, that is, the inability to isolate target sounds (Alcocer et al., 2022; Sugai et al., 2026). The application of an ad hoc designed classifier to measure anthropogenic noise is not new in ecological studies (Fairbrass et al., 2019). However, the cumulative nature of the SANE index, together with its integration of median amplitude, provides a finer description of noise levels and outperforms indices based on the proportion of time in which anthropogenic sounds are detected (e.g. CityAnthroNet activity level).

4.1 | SANE performance in real and simulated soundscapes

We found that frequency-based indices produced a counterintuitive result along the urbanization gradient, with both invNDSI and PSD (1–2 kHz) being higher in moderate than in high urbanization sites (Table S4). As suggested by the simulation results, these inconsistencies likely stem from the inability of the frequency-based indices to avoid the influence of non-target sounds, introducing natural noise into the estimation of anthropogenic sound levels. Biological sounds falling within the anthrophony frequency range, that is, 1–2 kHz (Pijanowski et al., 2011), for example, calls emitted

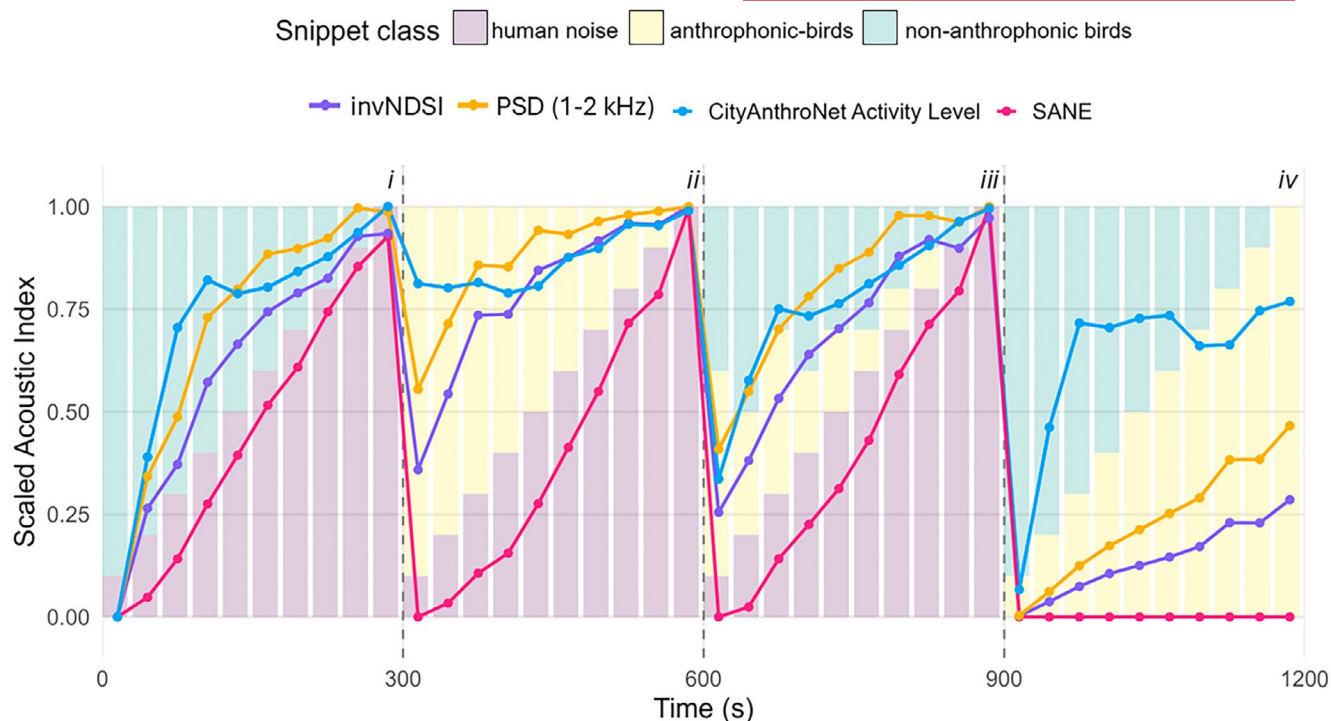


FIGURE 5 Temporal patterns of the invNDSI, PSD (1–2 kHz), CityAnthroNet Activity Level and global SANE, averaged across the 100 simulated soundscapes. All index values were scaled to the [0–1] range for direct comparison. Points represent the mean value of each index for a 30-second time step; lines connect consecutive time steps to show temporal trends. Background bars represent the proportion of 3-second snippets constituting each 30-second time step pertaining to a specific class. Simulated soundscapes consisted of four sequential phases delimited in the figure by the dashed lines: (i) only non-anthropophonic birds and human noise; (ii) only anthropophonic birds and human noise; (iii) comprising anthropophonic and non-anthropophonic birds and human noise; (iv) only anthropophonic and non-anthropophonic birds.

by the Eurasian Wild Boar or by the Yellow-legged Gull can make undisturbed areas be recognized as acoustically polluted. Similarly, several sources of anthropogenic noise, such as car horns or high-revving engines, that contribute significantly to the overall human acoustic contribution to the soundscape, can express noticeable acoustic energy within the theoretical biophonic frequency range (2–11 kHz).

The lower invNDSI and PSD values observed in the early morning (Figure 3a) can be explained by the presence of nocturnal non-target sounds, like insects' stridulation. In contrast, the CityAnthroNet activity increased sharply in the morning, followed by a decline in intensity, across the three urbanization classes. Such temporal patterns are likely associated with the acoustic activity of bird choruses, which typically dominate the soundscape (Ross et al., 2018; Staicer et al., 2019) and might be partially identified as anthropogenic sounds. Unlike the other indices, the global SANE consistently showed a progressive increase from dawn onwards, continuing until the end of the recording period across all urbanization classes. An important difference between the SANE and the CityAnthroNet activity level can be clearly noticed from their pattern in the high urbanization cluster (Figure 3). In the morning hours, both indices tend to rise progressively, but the CityAnthroNet activity level reaches a plateau. This is ascribable to its mathematical formulation: when every second of a recording contains human noises that are recognized with high confidence, CityAnthroNet saturates and stabilizes

to 1. Conversely, because the SANE index accounts for sound intensity, it does not reach an upper limit.

All indices were confused by the presence of anthropophonic birds, except for the SANE, thanks to the BirdNET ability of discriminating between sound classes. Minimum values of frequency-based indices were generally higher in the presence of anthropophonic bird sounds (Figure 5), and both invNDSI and PSD (1–2 kHz) and CityAnthroNet activity level increased with the increase of anthropophonic birds (Figure 5ii–iv). While this was expected for frequency-based indices, it was not for CityAnthroNet. Yet, this might be explained considering this classifier's training dataset, in which anthropophony represented more than 80% of sounds (Fairbrass et al., 2019). An imbalance in the training class sample size can produce a classifier biased towards the target class that flags inputs that only vaguely resemble the training data. Double-class classifiers, such as CityAnthroNet, are particularly sensitive to this issue (Wei & Dunbrack, 2013; Zheng & Jin, 2020). The SANE index was unique in remaining flat in the fourth time block of the simulated soundscapes (Figure 5iv), recognizing the complete absence of target sounds, thus proving its robustness to this issue.

SANE was also the most consistent across all simulated soundscapes in describing the level of human noise. Frequency-based indices, such as NDSI and 1–2 kHz PSD, necessitate hundreds of hours of recording before their values converge on a representative measure of the soundscape (Bradfer-Lawrence et al., 2019). As a result,

their values on a single audio recording are often unreliable, as demonstrated by the high variability across simulated soundscapes (Figure S2a,b). In contrast, the SANE was able to quantify disturbance at a very fine temporal scale, suggesting reliable application even at the single-recording level.

4.2 | SANE decomposability

Another substantial advantage of the proposed approach is the possibility of characterizing anthropogenic soundscapes by isolating different classes of anthropogenic sounds. Considering that different sources of noise can prompt responses of different nature and intensity in wild fauna (Francis & Barber, 2013; Shannon et al., 2015), the decomposability of the SANE index offers a considerable advantage over other indices. Depending on the focus of their investigations, researchers may be interested in isolating only a specific facet of the anthrophony or assessing the diverse impacts of different human-related noises. For instance, Randler (2006) have shown that dog barks are perceived as an immediate predatory risk by the Eurasian Coot, while McClure et al. (2013) have shown that motor vehicles are perceived as background noise by several bird species, causing acoustic signals masking.

Furthermore, thanks to the decomposability, it is possible to determine which are the main components of the global SANE, and therefore which disturbance classes contribute the most to the level of human noise (Figure 4). In our case study, for instance, we were able to determine that the main contributors to the global SANE values varied depending on the level of urbanization, with the Engine class being dominant in both low and moderate urbanized areas, and the Siren class contributing the most in highly urbanized sites.

4.3 | AI-classifier updates outcomes

AI classifiers are going to improve their performance in the future, which will inevitably influence the absolute values of the SANE index. Current classifiers, such as BirdNET, achieved almost perfect precision for anthropogenic sounds, but future updates could likely enhance detection sensitivity and expand the range of identifiable sound categories (i.e. recall of anthropogenic sounds), thereby limiting direct comparability of absolute SANE values between studies conducted with different classifier versions (or classifiers). Nonetheless, relative comparisons remain meaningful, as sites identified as more affected by anthropogenic noise are logically expected to be consistently recognized as such, regardless of classifier versions given sufficient sampling effort. Moreover, unlike traditional indices (Sugai et al., 2026), the SANE index performance is bound to increase as classifiers reliability continues to increase, as quantifying the acoustic energy associated with classified anthropogenic sounds provides a complementary, and ecologically particularly meaningful, information, as varying intensity and type of anthropogenic noises can induce responses of different nature

and magnitude in wildlife (Engel et al., 2024). Nonetheless, we recommend caution when directly comparing absolute SANE values across studies that differ in classifier type, version, or sampling design.

4.4 | Potential limits

The computation of the SANE index is relatively fast; however, the preparation of the dataset can be time-consuming. This is because sound-based AI classifiers can be computationally intensive. For the human-related noise classes for which we were able to compute performance metrics, precision was very high, while recall scores were comparatively lower. This is particularly important in describing ecological patterns, where false positives are often more undesirable than false negatives (Clare et al., 2020; Guillera-Aroita et al., 2017). Thus, although some signals were inevitably missed, this loss did not distort soundscape patterns, as the high precision allowed us to reliably capture the temporal dynamics of the four retained classes.

It is important to consider that evidence suggests a less effective performance of acoustic indices in soundscapes characterized by high species richness (Pan et al., 2024); it is therefore possible that AI classification performance may vary with soundscape complexity, thus influencing the SANE values. Our real-world case study represents a relatively heterogeneous and acoustically complex environment, spanning from relatively biodiverse areas to poor and highly disturbed areas (Russo et al., 2025), yet it remains representative of a temperate environment. Classifier performance might decrease in highly rich areas (e.g. tropical humid areas), potentially limiting comparability among habitats. While the validation process enables researchers to avoid the inclusion of non-target sounds by maximizing precision, we highly recommend caution when interpreting results across countries. A future validation of SANE in highly biodiverse ecosystems can elucidate the comparability of index values in very different ecosystems.

4.5 | Recommendations

We recommend that authors using the SANE explicitly report sampling design, including microphones' characteristics and settings, and the model and version of the AI classifier used to allow direct comparability.

Furthermore, the potential influence of geophonic sounds should be carefully considered when applying SANE. While low-intensity geophonies (e.g. flowing water or light wind), such as those included in our dataset, can be considered part of the background soundscape and are unlikely to substantially bias results, recordings dominated by intense meteorological events (e.g. heavy rain or strong wind) should be avoided or excluded. Such conditions are known to mask target acoustic signals and compromise the acoustic analyses (Bradfer-Lawrence et al., 2019; Sánchez-Giraldo et al., 2020). This

practice is consistent with standard approaches in ecoacoustics (Metcalf et al., 2020) and should be applied whether the focus is biophony or anthrophony.

In addition, given that SANE is based on acoustic amplitude, particular attention should be paid to recorder placement. Very loud sound sources occurring near the microphone may disproportionately influence the index and lead to misleading interpretations. Although the Median Amplitude index is somewhat robust to outliers, thus partially solving this issue, we still recommend designing sampling schemes that minimize placing recorders near predictable or constant noise sources to ensure more representative and robust measurements of the acoustic environment.

5 | CONCLUSION

AI-classifiers are powerful tools with high potential for bioacoustic research, but they are only one part of the ecological research process (Sugai et al., 2026). Here, we proposed a synergy between artificial intelligence and acoustic indices, through the novel SANE index, which has been shown to improve the quality of anthropogenic noise quantification compared to CityAnthroNet and traditional acoustic indices. This improvement stems from SANE's ability to not simply estimate the quantity of anthropogenic sounds present in the soundscape, an objective already achievable with AI-based classifiers, but also to quantify the magnitude of noise exposure.

By leveraging an AI-based classifier's ability to detect multiple human-made sound classes, SANE captures the intensity of anthropogenic noise, and it has the potential to assess the relative contributions of different noise types to the overall acoustic pollution. Knowing the nature of the source of anthropogenic noise can enhance ecological interpretation of its effects, as different anthropogenic sound sources can bring about different behavioural and physiological responses (Engel et al., 2024). Furthermore, not being tied to a specific frequency range, SANE can also be computed only in a specific spectral window, like the human-audible range or the ultrasonic one. This SANE's characteristic could open new research avenues, for example, the study of the effect of ultrasonic anthropogenic noise on bats that remains unexplored (Li et al., 2025).

Here we have applied the SANE index on the BirdNET classifier, but the SANE workflow can be easily adapted to different classifiers (e.g. a custom-trained classifier for anthropogenic sounds). With the release of updated versions of BirdNET or new classification algorithms, its accuracy in depicting levels of anthropogenic acoustic disturbance can further improve in the future. Thus, irrespective of the classifier used, this work emphasizes the importance of integrating artificial intelligence tools with acoustic indices to better assess the intensity and the nature of human noises, offering a reliable and flexible solution for acoustical research.

AUTHOR CONTRIBUTIONS

Matteo Giuliani, Luca Francesco Russo, Davide Mirante, Andrea Zampetti and Luca Santini conceived the ideas and designed

methodology; Matteo Giuliani collected the data; Matteo Giuliani, Luca Francesco Russo and Davide Mirante analysed the data; Matteo Giuliani and Luca Santini led the writing of the manuscript. All authors contributed critically to the drafts and gave final approval for publication.

ACKNOWLEDGEMENTS

MG, LFR and LS were supported by the PRIN Urbis grant (B53D23012400001) funded by the Italian Ministry for Universities and Research (MUR).

CONFLICT OF INTEREST STATEMENT

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

PEER REVIEW

The peer review history for this article is available at <https://www.webofscience.com/api/gateway/wos/peer-review/10.1111/2041-210x.70344>.

DATA AVAILABILITY STATEMENT

Tutorial, functions and part of the data are available via <https://doi.org/10.5281/zenodo.17242301> (Giuliani et al., 2025). The function to calculate the SANE index is freely accessible on the GitHub repository at <https://github.com/matpagle/sane>. Moreover, users can directly import the function into the environment by calling source (<https://github.com/matpagle/sane/blob/main/sanefunction.R>) within R.

STATEMENT ON INCLUSION

Our study was conducted by a research team composed entirely of scientists from the country where the study took place, all of whom share similar demographic characteristics. We acknowledge that this lack of diversity may have limited the range of perspectives represented in our work. We plan to address this limitation in future research by fostering more inclusive and diverse collaborations.

ORCID

Matteo Giuliani  <https://orcid.org/0009-0008-5959-0005>

Davide Mirante  <https://orcid.org/0009-0003-2258-0907>

Luca Francesco Russo  <https://orcid.org/0000-0003-4379-6160>

Andrea Zampetti  <https://orcid.org/0009-0001-0495-5416>

Luca Santini  <https://orcid.org/0000-0002-5418-3688>

REFERENCES

- Alcocer, I., Lima, H., Sugai, L. S. M., & Llusia, D. (2022). Acoustic indices as proxies for biodiversity: A meta-analysis. *Biological Reviews*, 97(6), 2209–2236. <https://doi.org/10.1111/brv.12890>
- Bradfer-Lawrence, T., Gardner, N., Bunnefeld, L., Bunnefeld, N., Willis, S. G., & Dent, D. H. (2019). Guidelines for the use of acoustic indices in environmental research. *Methods in Ecology and Evolution*, 10(10), 1796–1807. <https://doi.org/10.1111/2041-210X.13254>
- Brown, C. L., Hardy, A. R., Barber, J. R., Fristrup, K. M., Crooks, K. R., & Angeloni, L. M. (2012). The effect of human activities and their

- associated noise on ungulate behavior. *PLoS One*, 7(7), e40505. <https://doi.org/10.1371/journal.pone.0040505>
- Cignini, B., & Zapparoli, M. (1996). *Atlante degli uccelli nidificanti a Roma*. Fratelli Palombi Editore.
- Clare, J., Townsend, P., & Zuckerberg, B. (2020). Generalized model-based solutions to false positive error in species detection/non-detection data. *Ecology*, 101(11), e03241. <https://doi.org/10.1002/ecy.3241>
- Collins, A. C., Vickers, T. W., & Shilling, F. M. (2022). Behavioral responses to anthropogenic noise at highways vary across temporal scales. *Frontiers in Ecology and Evolution*, 10, 891595. <https://doi.org/10.3389/fevo.2022.891595>
- Cruz, A., Kleist, N., Guralnick, R., Francis, C., & Lowry, C. (2018). Chronic anthropogenic noise disrupts glucocorticoid signaling and has multiple effects on fitness in an avian community. *Proceedings of the National Academy of Sciences of the United States of America*, 115(3), E648–E657. <https://doi.org/10.1073/pnas.1709200115>
- Depraetere, M., Pavoine, S., Jiguet, F., Gasc, A., Duval, S., & Sœur, J. (2012). Monitoring animal diversity using acoustic indices: Implementation in a temperate woodland. *Ecological Indicators*, 13(1), 46–54. <https://doi.org/10.1016/j.ecolind.2011.05.006>
- Engel, M. S., Young, R. J., Davies, W. J., Waddington, D., & Wood, M. D. (2024). A systematic review of anthropogenic noise impact on avian species. *Current Pollution Reports*, 10, 684–709. <https://doi.org/10.1007/s40726-024-00329-3>
- Fairbrass, A. J., Firman, M., Williams, C., Brostow, G. J., Titheridge, H., & Jones, K. E. (2019). CityNet—Deep learning tools for urban eco-acoustic assessment. *Methods in Ecology and Evolution*, 10(2), 186–197. <https://doi.org/10.1111/2041-210X.13114>
- Francis, C. D., & Barber, J. R. (2013). A framework for understanding noise impacts on wildlife: An urgent conservation priority. *Frontiers in Ecology and the Environment*, 11(6), 305–313. <https://doi.org/10.1890/120183>
- Fuller, R. A., Warren, P. H., & Gaston, K. J. (2007). Daytime noise predicts nocturnal singing in urban robins. *Biology Letters*, 3(4), 368–370. <https://doi.org/10.1098/rsbl.2007.0134>
- Gemmeke, J. F., Ellis, D. P. W., Freedman, D., Jansen, A., Lawrence, W., Moore, R. C., Plakal, M., & Ritter, M. (2017). AudioSet: An ontology and human-labeled dataset for audio events. In ICASSP. IEEE.
- Giuliani, M., Mirante, D., Abbondanza, E., & Santini, L. (2024). Acoustic indices fail to represent different facets of biodiversity. *Ecological Indicators*, 166, 112451. <https://doi.org/10.1016/j.ecolind.2024.112451>
- Giuliani, M., Mirante, D., Russo, L. F., Zampetti, A., & Santini, L. (2025). SANE: An Index of Anthropogenic Noise Levels for Wildlife Research in Terrestrial Ecosystems. <https://doi.org/10.5281/zenodo.17242301>
- Gomes, L., Sousa-Lima, R., Solé, M., & Baumgarten, J. (2022). Influence of anthropogenic sounds on insect, anuran, and bird acoustic signals: A meta-analysis. *Frontiers in Ecology and Evolution*, 10, 827440. <https://doi.org/10.3389/fevo.2022.827440>
- Guillera-Aroita, G., Lahoz-Monfort, J., van Rooyen, A., Weeks, A., & Tingley, R. (2017). Dealing with false-positive and false-negative errors about species occurrence at multiple levels. *Methods in Ecology and Evolution*, 8(9), 1081–1091. <https://doi.org/10.1111/2041-210X.12743>
- Higham, V., Deal, N. D. S., Chan, Y. K., Chanin, C., Davine, E., Gibbings, G., Keating, R., Kennedy, M., Reilly, N., Symons, T., Vran, K., & Chapple, D. G. (2021). Traffic noise drives an immediate increase in call pitch in an urban frog. *Journal of Zoology*, 313, 307–315. <https://doi.org/10.1111/jzo.12866>
- Hildebrand, J. A. (2009). Anthropogenic and natural sources of ambient noise in the ocean. *Marine Ecology Progress Series*, 395, 5–20. <https://doi.org/10.3354/meps08353>
- Hunter, J. D. (2007). Matplotlib: A 2D graphics environment. *Computing in Science & Engineering*, 9(3), 90–95. <https://doi.org/10.1109/MCSE.2007.55>
- Jacewicz, E., Alexander, J. M., & Fox, R. A. (2023). Introduction to the special issue on perception and production of sounds in the high-frequency range of human speech. *The Journal of the Acoustical Society of America*, 154(5), 3168–3172.
- Kahl, S., Wood, C. M., Eibl, M., & Klinck, H. (2021). BirdNET: A deep learning solution for avian diversity monitoring. *Ecological Informatics*, 61, 101236. <https://doi.org/10.1016/j.ecoinf.2021.101236>
- Kaiser, K., & Hammers, J. L. (2009). The effect of anthropogenic noise on male advertisement call rate in the neotropical treefrog *Dendropsophus triangulum*. *Behaviour*, 146(8), 1053–1069. <http://www.jstor.org/stable/40296115>
- Kasten, E. P., Gage, S. H., Fox, J., & Joo, W. (2012). The remote environmental assessment laboratory's acoustic library: An archive for studying soundscape ecology. *Ecological Informatics*, 12, 50–67. <https://doi.org/10.1016/j.ecoinf.2012.08.001>
- Lampe, U., Reinhold, K., & Schmoll, T. (2014). How grasshoppers respond to road noise: Developmental plasticity and population differentiation in acoustic signalling. *Functional Ecology*, 28(3), 660–668. <https://doi.org/10.1111/1365-2435.12215>
- Li, A., Tian, X., Ma, N., & Luo, J. (2025). Habitat preference contributes to explaining the varied sensitivity of bats to anthropogenic noise. *Biological Conservation*, 302, 110974. <https://doi.org/10.1016/j.biocon.2025.110974>
- Ligges, U., Krey, S., Mersmann, O., & Schnackenberg, S. (2023). tuneR: Analysis of music and speech. R package. <https://CRAN.R-project.org/package=tuneR>
- Loss, S., Hovick, T., & Duquette, C. (2021). A meta-analysis of the influence of anthropogenic noise on terrestrial wildlife communication strategies. *Journal of Applied Ecology*, 58(5), 1113–1125. <https://doi.org/10.1111/1365-2664.13880>
- McClure, C. J., Ware, H. E., Carlisle, J., Kaltenecker, G., & Barber, J. R. (2013). An experimental investigation into the effects of traffic noise on distributions of birds: Avoiding the phantom road. *Proceedings of the Royal Society B*, 280(1773), 20132290. <https://doi.org/10.1098/rspb.2013.2290>
- Metcalfe, O. C., Lees, A. C., Barlow, J., Marsden, S. J., & Devenish, C. (2020). hardRain: An R package for quick, automated rainfall detection in ecoacoustic datasets using a threshold-based approach. *Ecological Indicators*, 109, 105793. <https://doi.org/10.1016/j.ecolind.2019.105793>
- Pan, W., Goodale, E., Jiang, A., & Mammides, C. (2024). The effect of latitude on the efficacy of acoustic indices to predict biodiversity: A meta-analysis. *Ecological Indicators*, 159, 111747. <https://doi.org/10.1016/j.ecolind.2024.111747>
- Paterson, J. E., Bortolotti, L. E., Kowal, P. D., Pidwerbesky, A. J., & Devries, J. H. (2024). Predicting the effects of land cover change on biodiversity in prairie Canada using species distribution models. *Biological Conservation*, 298, 110754. <https://doi.org/10.1016/j.biocon.2024.110754>
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, É. (2011). Scikit-learn: Machine learning in python. *Journal of Machine Learning Research*, 12, 2825–2830.
- Pérez-Granados, C., Funosas, D., Morant, J., Marín Gómez, O. H., Mendoza, I., Mohedano-Munoz, M. A., Santamaría, E., Bastianelli, G., Márquez-Rodríguez, A., Budka, M., Bota, G., De la Peña-Rubio, J. M., De La Morena, E. G., Santa-Cruz, M., De la Nava, P., Fernández-Tizón, M., Sánchez-Mateos, H., Barrero, A., Traba, J., ... Sebastián-González, E. (2025). Optimisation of passive acoustic bird surveys: A global assessment of BirdNET settings. Research Square Preprint <https://doi.org/10.21203/rs.3.rs-6633549/v1>
- Pijanowski, B. C., Villanueva-Rivera, L. J., Dumyahn, S. L., Farina, A., Krause, B. L., Napoletano, B. M., Gage, S. H., & Pieretti, N. (2011). Soundscape ecology: The science of sound in the landscape. *Bioscience*, 61(3), 203–216. <https://doi.org/10.1525/bio.2011.61.3.6>

- Plat, N., Moreews, L., Schillé, L., Rivoal, J. B., & Jactel, H. (2025). Conservation biological control in the forest: A case study with the pine processionary moth. *Biological Control*, 202, 105702. <https://doi.org/10.1016/j.biocontrol.2025.105702>
- Pollock, L. J., Kitzes, J., Beery, S., Gaynor, K. M., Jarzyna, M. A., Mac Aodha, O., Meyer, B., Rolnick, D., Taylor, G. W., Tuia, D., & Berger-Wolf, T. (2025). Harnessing artificial intelligence to fill global shortfalls in biodiversity knowledge. *Nature Reviews Biodiversity*, 1, 166–182. <https://doi.org/10.1038/s44358-025-00022-3>
- Putland, R. L., Constantine, R., & Radford, C. A. (2017). Exploring spatial and temporal trends in the soundscape of an ecologically significant embayment. *Scientific Reports*, 7, 5713. <https://doi.org/10.1038/s41598-017-06347-0>
- Python Software Foundation. (2016). *Python (Version 2.7.12) [Computer software]*. Python Software Foundation. <https://www.python.org/>
- R Core Team. (2024). *R: A language and environment for statistical computing [Computer software]*. R Foundation for Statistical Computing. <https://www.R-project.org/>
- Randler, C. (2006). Disturbances by dog barking increase vigilance in coots (*Fulica atra*). *European Journal of Wildlife Research*, 52(4), 265–270. <https://doi.org/10.1007/s10344-006-0049-z>
- Ross, S. R. P. J., Friedman, N. R., Dudley, K. L., Yoshimura, M., Yoshida, T., & Economo, E. P. (2018). Listening to ecosystems: Data-rich acoustic monitoring through landscape scale sensor networks. *Ecological Research*, 33(1), 135–147. <https://doi.org/10.1007/s11284-017-1509-5>
- Russo, L. F., Andreello, M., Giuliani, M., Ancillotto, L., Carboni, M., La Bella, G., Martelli, F., & Santini, L. (2025). Mapping the urban landscape at multiple ecological scales. *Urban Forestry & Urban Greening*, 110, 128849. <https://doi.org/10.1016/j.ufug.2025.128849>
- Sánchez-Giraldo, C., Bedoya, C. L., Morán-Vásquez, R. A., Isaza, C. V., & Daza, J. M. (2020). Ecoacoustics in the rain: Understanding acoustic indices under the most common geophonic source in tropical rainforests. *Remote Sensing in Ecology and Conservation*, 6(3), 248–261. <https://doi.org/10.1002/rse2.162>
- Shannon, G., McKenna, M. F., Angeloni, L. M., Crooks, K. R., Frstrup, K. M., Brown, E., Warner, K. A., Nelson, M. D., White, C., Briggs, J., McFarland, S., & Wittemyer, G. (2015). A synthesis of two decades of research documenting the effects of noise on wildlife. *Biological Reviews*, 91(4), 982–1005. <https://doi.org/10.1111/brev.12207>
- Slabbekoorn, H., & Peet, M. (2003). Birds sing at a higher pitch in urban noise. *Nature*, 424(6946), 267. <https://doi.org/10.1038/424267a>
- Staicer, C. A., Spector, D. A., & Horn, A. G. (2019). The dawn chorus and other diel patterns in acoustic signaling. In *Ecology and evolution of acoustic communication in birds* (pp. 426–453). Cornell University Press.
- Stowell, D., & Plumbley, M. D. (2013). *An open dataset for research on audio field recording archives: freefield1010*. arXiv:1309.5275/AES Conference Proceedings, 2014.
- Sueur, J., Aubin, T., & Simonis, C. (2008). Seewave, a free modular tool for sound analysis and synthesis. *Bioacoustics*, 18(2), 213–226. <https://doi.org/10.1080/09524622.2008.9753600>
- Sueur, J., Farina, A., Gasc, A., Pieretti, N., & Pavoine, S. (2008). Acoustic indices for biodiversity assessment and landscape investigation. *Acta Acustica united with Acustica*, 100(4), 772–781. <https://doi.org/10.3813/AAA.918757>
- Sugai, L. S. M., Balantic, C., Clink, D. J., Ramesh, V., Kahl, S., Klinck, H., & Wood, C. M. (2026). Acoustic indices are not useful for biodiversity research. *Methods in Ecology and Evolution*, 17(5), 1506–1518. Portico. <https://doi.org/10.1111/2041-210x.70285>
- Sweet, K. A., Sweet, B. P., Gomes, D. G. E., Francis, C. D., & Barber, J. R. (2022). Natural and anthropogenic noise increase vigilance and decrease foraging behaviors in song sparrows. *Behavioral Ecology*, 33(1), 288–297. <https://doi.org/10.1093/beheco/arab141>
- Wei, Q., & Dunbrack, R. (2013). The role of balanced training and testing data sets for binary classifiers in bioinformatics. *PLoS One*, 8(7), e67863. <https://doi.org/10.1371/journal.pone.0067863>
- Welch, P. D. (1967). The use of fast Fourier transform for the estimation of power spectra: A method based on time averaging over short, modified periodograms. *IEEE Transactions on Audio and Electroacoustics*, 15(2), 70–73. <https://doi.org/10.1109/TAU.1967.1161901>
- Zheng, W., & Jin, M. (2020). The effects of class imbalance and training data size on classifier learning: An empirical study. *SN Computer Science*, 1(1), 74. <https://doi.org/10.1007/s42979-020-0074-0>

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

Figure S1. Map showing the distribution of the sampling points of this study.

Table S1. Complete list of snippets used for simulated soundscapes creation.

Figure S2. Variability of acoustic indices across simulated soundscapes.

Table S2. List of Pearson correlation coefficients among acoustic indices computed on real soundscape for each urbanization category.

Table S3. List of Pearson correlation coefficients among acoustic indices computed on simulated soundscape for each phase: (i) only non-anthropogenic birds and human noise; (ii) only anthropogenic birds and human noise; (iii) comprising anthropogenic and non-anthropogenic birds, and human noise; (iv) only anthropogenic and non-anthropogenic birds.

Table S4. Results of Mann-Whitney *U* tests with Bonferroni correction applied separately for each acoustic index.

How to cite this article: Giuliani, M., Mirante, D., Russo, L. F., Zampetti, A., & Santini, L. (2026). SANE: An index of anthropogenic noise levels for wildlife research in terrestrial ecosystems. *Methods in Ecology and Evolution*, 00, 1–13. <https://doi.org/10.1111/2041-210x.70344>