Unsupervised analysis of background noise sources in active offices

Published Version:
De Salvio D., D'Orazio D., Garai M. (2021). Unsupervised analysis of background noise sources in active offices. THE JOURNAL OF THE ACOUSTICAL SOCIETY OF AMERICA, 149(6), 4049-4060 [10.1121/10.0005129].

Availability:
This version is available at: https://hdl.handle.net/11585/835569 since: 2024-09-06

Published:
DOI: http://doi.org/10.1121/10.0005129

Unsupervised analysis of background noise sources in active offices

Domenico De Salvio,[1] Dario D'Orazio,[1] and Massimo Garai[1]

Department of Industrial Engineering (DIN), University of Bologna,

Viale Risorgimento 2, 40136 Bologna, Italy[a]

Inside open-plan offices, background noise affects the workers' comfort, influencing their productivity. Recent approaches identify three main source categories: mechanical sources (HVAC equipment, office devices, etc.), outdoor traffic noise, and human sources (speech). While the first two groups are taken into account by technical specifications, human noise is still often neglected. The present paper proposes two procedures to identify the human and mechanical noise sources during working hours, based on machine-learning techniques. Two unsupervised clustering methods, specifically Gaussian Mixture Model and K-means, were used to separate the recorded sound pressure levels recorded finding the candidate models. Thus, the clustering validation was used to find the number of sound sources within the office and then, statistical and metrical features were used to label the sources. The results were compared with the common parameters used in noise monitoring in offices, i.e. the equivalent continuous and the 90th percentile levels. The spectra obtained by the two Algorithms match with the expected shapes of human speech and mechanical noise tendencies. The outcomes validate the robustness and reliability of these procedures.

[a]dario.dorazio@unibo.it

I. INTRODUCTION

Different kinds of noise can affect the individual perception, depending on personal factors (Ellermeier *et al.*, 2001), on the task to do (Braat-Eggen *et al.*, 2019), and on the nature of the noise source (Koskela *et al.*, 2014). To ensure a more comfortable workspace, it is important to take into account which work tasks are performed, the spatial distribution of the workstations and the balance between ease of communication and concentration (Perrin Jegen and Chevret, 2017). In many cases, distraction in the work place is not strictly related to the sound pressure level but more to the effect of irrelevant sound (Ellermeier and Zimmer, 2014; Jones, 1999). In this latter case, the most distracting noise is due to speech and the unintentional listening of the workmates conversations (Braat-Eggen *et al.*, 2019; Hongisto, 2005). The perception of mechanical noise has been investigated by classifying its characteristics (Iannace *et al.*, 2018). In some cases, this kind of noise sources could even increase performance (Alimohammadi and Ebrahimi, 2017). These instances are faced by acoustic consultants, who should balance the acoustic absorption inside open plan offices while keeping the background noise at a reasonable level, because it is fundamental to guarantee a low speech intelligibility (Di Blasio *et al.*, 2019; Schlittmeier and Liebl, 2015).

The spatial distribution of Speech Transmission Index (STI) was proved to be a workers' performance metrics (Hongisto, 2005). The STI depends on the acoustical quality of the room and on the background noise, even if international regulations on open-plan office acoustic quality (ISO 3382, 2012) state that STI must be evaluated by neglecting the contribution of human noise. This condition that tends to underestimate the real acoustic environment (Harvie-Clark *et al.*, 2019). It implies the assumption that the most distracting situation is due to a single talker and not to a multi-talker scenario (Yadav and Cabrera, 2019; Yadav *et al.*, 2017). It is based on the fact that measurements in an unoccupied condition are much more easily performed, when the role of the human activities involved in background noise can be neglected. Consequently, it is neglected also during the evaluation of the acoustic parameters. Open plan offices constitute dynamic scenarios in which people are no longer to be considered only as sensitive receivers but also as sound sources themselves (Renz *et al.*, 2018a,b). A more specific background noise condition can be selected to post-processing the STI values, so, in light of this, the measurement of this parameter is crucial (D'Orazio *et al.*, 2018; Rindel, 2018). Different criteria have been proposed with the aim to produce an objective descriptor of an acoustic environment, that enables people to estimate its impact on their comfort and productivity (Renz *et al.*, 2019; Vellenga *et al.*, 2017).

In the field of room acoustics, data-analysis based techniques (Bianco *et al.*, 2019) were used to measure the background noise due to human activity into classrooms (D'Orazio *et al.*, 2020; Hodgson *et al.*, 1999) and - in a preliminary way - in open plan offices too (Dehlbæk *et al.*, 2016).

The aim of this work is to identify the sound sources via an unsupervised statistical analysis of long-term monitoring, including the number of sources, their origin and the sound pressure level they produce. The data population obtained from the recording done with a sound level meter can be processed with algorithms used in unsupervised learning to find pattern and create clusters. Then, each sound source can be reliably associated to each cluster. In particular, machine learning techniques can help in precisely delineating human noise. The results of the algorithms are compared with standard procedures to measure background noise.

II. SOUND SOURCE DETECTION THROUGH DATA ANALYSIS

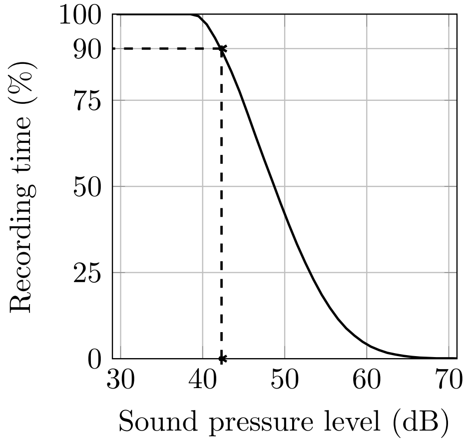
In the field of open-plan office acoustics, noise monitoring is often made with percentile levels. Several thresholds of percentiles are used, often compared with the equivalent continuous level L_{eq}. Vellenga-Persoon et al. used L_5 in the so called "liveliness ratio" (Vellenga *et al.*, 2017), whereas percentiles L_5, L_{10}, L_{90}, and L_{95} were compared by Renz et al. (Renz *et al.*, 2019).

This kind of approach requires either the knowledge of the distribution of the occurrences of the sound levels in the environment during monitoring time, or a supervision by the operator. For these reasons, unsupervised algorithms can represent a useful tool for accurate monitoring without the need of human supervision.

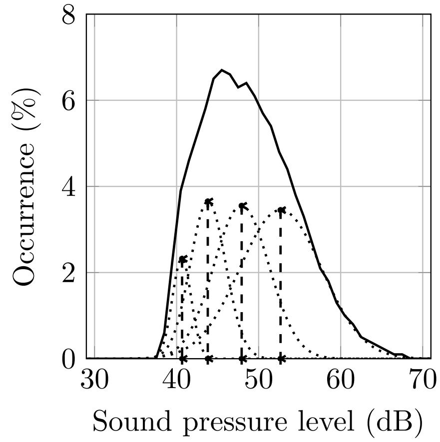
A. Clustering algorithms

Clustering algorithms allow to identify different candidate noise sources by analyzing the data collected from a recording. In this section, the Gaussian Mixture Model (GMM) and the K-Means Clustering (KM) are introduced.

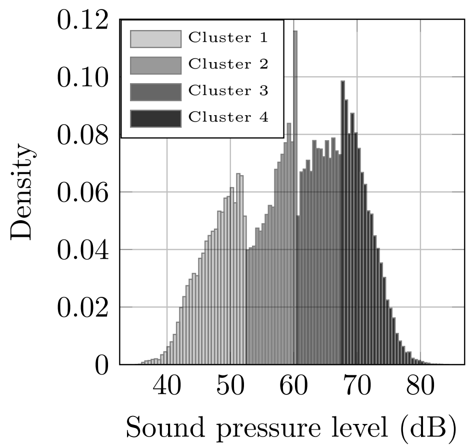
GMM is a clustering method which decomposes the original model data in a sum of gaussian curves. Assuming a set of observations $x_1, \ldots, x_n$ (e.g. the short-time equivalent levels recorded), the Gaussian probability density function $f(x_i)$ of these observations – in



(a) Percentile levels



(b) Gaussian Mixture Model



(c) K-means clustering

FIG. 1. The three unsupervised methods used in this work. In figure (a) the continuous line represent the cumulative distribution of the recorded SPL of a sound level meter and the $*$ corresponds to the 90th Percentile Level L_{90}. In figure (b) the continuous line represents the occurrences distribution of the same measurement. The asymmetrical distribution can be decomposed in four Gaussian curves. The mean values of Gaussian curves indicated with $*$ correspond to the sound levels attributed to each sound source. In figure (c) the four histograms represent four different clusters obtained via K-means clustering.

the following called *target density* – can be expressed as a sum of K Gaussian densities $f_k(x_i, \mu_k, \sigma_k^2)$:

$$f(x_i) \simeq \sum_{k=1}^{K} \pi_k f_k(x_i, \mu_k, \sigma_k^2) \tag{1}$$

where π_k are the so called *mixing proportions* (McLachlan, G.J. and Peel, D., 2000), non-negative quantities that sum to one; that is,

$$0 \leq \pi_k \leq 1 \quad (k = 1, \ldots, K)$$

and

$$\sum_{k=1}^{K} \pi_k = 1.$$

The likelihood function for a mixture model with K univariate Normal components is:

$$\mathcal{L}(x) = \prod_{i=1}^{n} \sum_{k=1}^{K} \pi_k f_k(x_i) = \prod_{i=1}^{n} \sum_{k=1}^{K} \pi_k \frac{1}{\sqrt{2\pi\sigma_k^2}} e^{-\frac{(x_i-\mu_k)^2}{2\sigma_k^2}}. \tag{2}$$

The equality in 1 is usually realized by maximum likelihood optimization algorithm, e.g. the Expectation-Maximization (EM) (Dempster *et al.*, 1977). In the context of background noise in open-plan offices, Dehlbaek et al. (Dehlbæk *et al.*, 2016) proposed a preliminary analysis based on GMM. The probability distribution function of equivalent levels recorded in several offices is fitted with one or more Gaussian curves. The means a Gaussian curve is taken as the sound pressure levels of a sound source. If two normal curves are used, then the higher mean is identified as human activity and the lower one as the background noise in the office. The contribution of human activity is taken into account only if the 10th statistical percentile of the corresponding curve is greater than the background noise measured in unoccupied condition.

While GMM is based on statistical properties of the data population, KM optimizes a metric distance of each single point data to form clusters. The set of observations $x_1, \ldots, x_n$ can be clustered into a set of K clusters, $C = \{c_k; k = 1, \ldots, K\}$, where μ_k is the mean of cluster c_k. The squared Euclidean distances between μ_k and the points in cluster c_k is defined as:

$$J(c_k) = \sum_{x_i \in c_k} ||x_i - \mu_k||^2. \tag{3}$$

The goal of K-means is to minimize the sum of the squared Euclidean distances over all K clusters:

$$J(C) = \sum_{k=1}^{K} \sum_{x_i \in c_k} ||x_i - \mu_k||^2. \tag{4}$$

The process converges to a local minimum in two steps: first, the optimal partition for a given set of μ_k is found; then, the cluster centroids are computed once C is fixed (Lloyd, 1982). A K-means clustering was preliminary used by Wang and Brill to monitor the noise levels in occupied and unoccupied conditions in several K-12 classrooms (Wang *et al.*, 2020).

B. Clustering validation

Clustering algorithms may produce redundant results, i.e. a number of clusters greater than the number of actual sound sources. Indeed, the maximum likelihood principle results in selecting the highest possible dimension (Schwarz, 1978). The clustering validation allows to assess the best model among candidates through specific metrics. In this work the Akaike Information Criterion (AIC) (Akaike, 1974) and the silhouette method (Rousseeuw, 1987) have been used to assess the optimal number of clusters for, respectively, GMM and KM.

AIC provides an assessment based on a reward for the goodness of fit to help in choosing the best candidate as well as a penalization for the complexity of the model. Assuming k as the number of estimated parameters in the model and $\mathcal{L}$ the likelihood function defined above, the AIC is:

$$AIC = 2k - 2\ln\left[\mathcal{L}(x)\right]. \tag{5}$$

The first term of eq. 5 is the penalization of the complexity, whereas the second term concerns the goodness of the fit. Thus, the greater the likelihood the lower the AIC. It follows that the lowest value indicates the best model. Plotting the AIC obtained with different values of K, the elbow of the curve highlights the optimal number of clusters. More in detail, the AIC coefficient estimates the error caused by the loss of information due to the statistical modelling of the initial data (Rodríguez, 2005). Instead of other information criteria like the Bayesian information criterion (BIC), the AIC assumes that all the candidate models are wrong, i.e. none of them is the true model that generated the data. Thus, the BIC seeks the true model, which is the most probable, whereas the AIC seeks the wrong model with the lowest loss of information, which is the most predictive referred to the initial data. Moreover, it has been shown that AIC performs better than other information criteria when the models are non-nested, i.e. one model is not a particular case of another (Gabbay *et al.*, 2011). The AIC has been chosen in the present work since all the candidate models are assumed wrong and non-nested.

The silhouette method is a quantitative assessment of the degree of separation among the clusters. Assuming i as a data point in the cluster A_m, the mean distance between i and the other data points in the same cluster, is:

$$a(i) = \frac{1}{|A_m| - 1} \sum_{i,j \in A_m, i \neq j} d(i,j) \tag{6}$$

where $d(i,j)$ is the distance between i and j in the cluster A_m.

Thereafter, the mean dissimilarity of i to another cluster B_n is defined as the mean distance between i and the other points l in B_n. Thus:

$$b(i) = \min \frac{1}{|B_n|} \sum_{l \in B_n, l \neq i} d(i,l) \tag{7}$$

is the shortest mean distance between i and all the other points in the other clusters. Of course, this is possible only with a number of clusters $K > 1$. The cluster with the smallest mean dissimilarity is defined as "neighbor" and represents the second-best choice for i. The silhouette value $s(i)$ is defined as:

$$\begin{cases} 1 - a(i)/b(i) & \text{if } a(i) < b(i), \\ 0 & \text{if } a(i) = b(i), \\ b(i)/a(i) - 1 & \text{if } a(i) > b(i). \end{cases} \tag{8}$$

Thus $-1 \leq s(i) \leq 1$, which means that i is properly clustered if $s(k)$ is near 1, while it is wrongly clustered if $s(i)$ is near -1, whereas an $s(i)$ near 0 means that i can be assigned to either A or B. The silhouette values $s(i)$ expresses how each data point is well clustered. Hence, the mean of each silhouette value of clusters $\bar{s}(i)$ can be considered as a metric for the whole clustering process. The silhouette coefficient SC, then is defined as:

$$SC = \max_k \bar{s}(k) \tag{9}$$

where k is the number of clusters. The silhouette coefficient, as well as being one of the most well-known clustering validation indices, is assessed as very viable among different kinds of dataset (Jauhiainen and Kärkkäinen, 2017).

III. METHOD

A. Checking through 1-day recordings

In order to evaluate the performance of the two methods, an office with four workstations was chosen as test room.

The office is placed in a building away from the city center and road traffic. This means that the expected soundscape involves only two main noise sources: human activity and mechanical systems. Moreover, the room has an acoustically treated ceiling and thus, it can be considered as a "dead" environment. Hence, the goal is to identify these two sound sources during work hours. This test office, given the above mentioned simplifications, allows to monitor the appropriate noise sources: human activity of several people at the same time, mechanical noise due to air conditioning systems (which could be switched off during measurements) and other office devices. The workplace layout is made up of four workstations (ws A, ws B, ws C and ws D) and a meeting table. Despite the small size of the office, the meeting table is far enough from the workstations to allow the analysis of speech at the position number 4 (in figure 2).

Sound pressure levels monitoring was carried out throughout an entire working day, so to allow to record enough data (ISO 22955, 2020). A statistical data population was obtained by recording short-time equivalent levels, 100 ms integration time, octave-band filtered (125 Hz - 4000 Hz), for an amount of time long enough to validate the central limit theorem.

Thereafter, the collected data were processed through the procedures shown in the following section. Furthermore, equivalent continuous L_{eq}, and percentile levels (L_{90}, L_{50}, and L_{10}) were extracted in order to compare the results with previous studies on this topic (Renz *et al.*, 2019).

B. Description of the three-steps procedures

The unsupervised analysis proposed here is based on two Algorithms. Algorithm 1 performs the clustering via GMM and AIC. Algorithm 2 performs the clustering via KM and silhouette. Both Algorithms act in three steps:

1. Preliminary clustering analysis, finding several numbers of candidate noise sources.
2. Selection of the best candidate through clustering validation.
3. Final clustering analysis and association of each cluster to a noise contribute on the basis of statistical (Algorithm 1) or temporal (Algorithm 2) conditions.

In Algorithm 1, the first step performs the clustering via GMM. The procedure has been repeated with a variable number of clusters $k = 1, ..., 10$. The EM algorithm returns the mean μ_k, the standard deviation σ_k, and the mixing proportions π_k of the each Gaussian curve (see eqs. 1 and 2). In order to achieve meaningful results, EM algorithm is initialized by means of the components, the covariance matrices, and the mixing proportions. An option has been set in order to replicate the algorithm several times starting from different

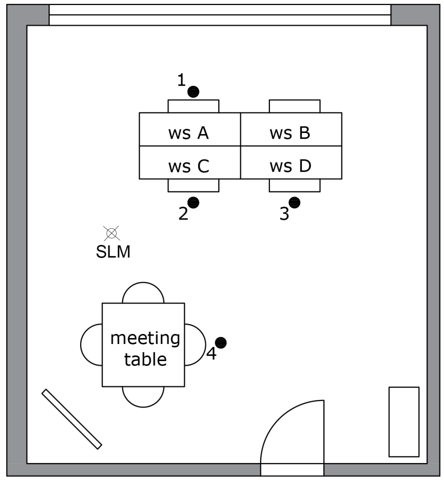


FIG. 2. Layout of the office with four workstations and a meeting table. The measurement positions (1 - 4) were used to evaluate the STI values and the role of irrelevant speech. The sound level meter, used for the long-time monitoring was placed between the workstations and the meeting table.

Algorithm 1: GMM and AIC

Input: x_i short-time levels octave-band filtered, $f(x_i)$ target distribution

Output: L_{human}; L_{mech}

init EM

init $L_{\text{mech}}, L_{\text{human}} = -\infty$

// first step

for $k = 1 : 10$ **do**

  $(\pi_k, \mu_k, \sigma_k) = \text{EM}(k, x_i)$

end

// second step

for $k = 1 : 10$ **do**

  $AIC(k) = 2k - 2\ln\left(\mathcal{L}(\pi_k, \mu_k, \sigma_k; f(x_i))\right)$

end

$E = \text{elbow}(AIC(k))$

// third step

$(\pi_j, \mu_j, \sigma_j) = \text{EM}(E, x_i)$

for $j = 1 : E$ **do**

  if *statistical condition* **then**

    $L_{\text{human}} = 10\log\left(10^{\mu_j/10} + 10^{L_{\text{human}}/10}\right)$

  end

  else

    $L_{\text{mech}} = 10\log\left(10^{\mu_j/10} + 10^{L_{\text{mech}}/10}\right)$

  end

end

Algorithm 2: KM and Silhouette

Input: x_i short-time levels octave-band filtered, $f(x_i)$ target distribution

Output: L_{human}; L_{mech}

init KM

init $L_{\text{mech}}, L_{\text{human}} = -\infty$

// first step

for $k = 1:10$ **do**

  $c_k = \text{KM}(k, x_i)$

end

// second step

for $k = 2:10$ **do**

  $s(k) = \text{Silhouette}(c_k; f(x_i))$

end

$A = k : \max_k (s(k))$

// third step

$c_j = \text{KM}(SC, x_i)$

for $j = 1 : SC$ **do**

  if *metrical condition* **then**

    $L_{\text{human}} = 10 \log \left(10^{\text{centroid}(c_j)/10} + 10^{L_{\text{human}}/10}\right)$

  end

  else

    $L_{\text{mech}} = 10 \log \left(10^{\text{centroid}(c_j)/10} + 10^{L_{\text{mech}}/10}\right)$

  end

end

points, then the maximum likelihood is fitted. A covariance matrix of diagonal type is set, whereas the mixing proportions are used with default parameters, which means that the initial values are uniform. In the second step of Algorithm 1, the optimal number of clusters is investigated through the AIC calculation according to equation 5. The goodness of fit is rewarded through the likelihood function and, at the same time, the model is penalized if it exceeds in complexity. The number k corresponding to the elbow of the curve is used to perform again the GMM with the optimal number of clusters (see figure 3). Then, the association among numerical and real sources existing within the office is made. Since in the dataset used in the present study the traffic noise is negligible, in the third step of Algorithm 1 the way to discern the type of source was statistical. In fact, the standard deviation is used as the parameter to distinguish the nature of the source, either mechanical or human. It is expected that a low s.d. belongs to the mechanical sources, whereas a high s.d. to human activity.

Instead, Algorithm 2 is based on KM and silhouette. The K-means clustering was set using the square Euclidean distance as the metric to be minimized within the cluster $\mathrm{c_k}$ and all over the k clusters (see eqs. 3 and 4). Then, as seen above for GMM, a specific option to replicate the algorithm starting from different points was set to avoid the use of the same centroids in the iterations. Also, Algorithm 2 was repeated with a variable number of clusters $k = 2, ..., 10$. Then, the silhouette method was used to choose the best model among candidates. The mean values of the silhouettes of each cluster provides a metric to evaluate the clustering goodness. Thus, the clustering validity is rated finding the highest silhouette coefficient SC (equation 9) among candidate models, which means for various number of clusters k, as described in the previous section. Then, KM is performed again with the optimal k. The subdivision between mechanical and human noise is based on a metrical hypothesis. The average distance of data points and the centroid within a cluster describes the density of clusters. A short distance can be associated to a mechanical source whereas a large value to human activity. The size of clusters can bring information about the frequency – in the temporal meaning – of the sound sources. For instance, a quiet office, with a low human activity within, will have a corresponding cluster with a large percentage of samples relative to the whole population.

C. Influence on Privacy-criteria

In order to evaluate the influence of background noise on the intelligibility of speech, a numerical model of the office was done using a ray tracing software to provide the modulation transfer function of the room. The model was calibrated using the reverberation time measured in the office, assuming that the diffusivity conditions are met. This means that the calibration is achieved when the difference between the measured and simulated reverberation times lies within the just noticeable difference (JND), which is ±5%, following the recommendations of the state-of-art (Vorländer, 2020). Once the model is set, the modulation transfer functions matrix (mtf) can be calculated from the simulated impulse response h(t) for each k-th octave band and the modulation frequency $\mathrm{f_m}$ (IEC 60268, 2020):

$$m_k(f_m) = \frac{\left|\int_0^\infty h_k^2(t)e^{-j2\pi f_m t}\mathrm{d}t\right|}{\int_0^\infty h_k^2(t)\mathrm{d}t}\left[1+10^{-\frac{\left(L_{S,k}-L_{N,k}\right)}{10}}\right]^{-1} \tag{10}$$

where $L_{S,k}$ is the useful speech level, which depends on the distance, and $L_{N,k}$ the background noise. Through this matrix, the Speech Transmission Index can be calculated for each workstation varying the background noise levels. Then, a predictive analysis of privacy criteria within the active office can be carried out as function of the sound sources extracted. This is very useful since it is directly linked to the studies about annoyance during work hours (Ebissou *et al.*, 2015). Four positions were chosen as representative of the activity performed inside the office: three near the workstations and one near the meeting table. According to ISO 3382-3, the simulation of privacy criteria was carried out using a directional sound source set at "normal" vocal effort. Different source-receiver configurations were simulated by moving the source in one position (from n.1 to n.4) and the receivers in the others. In order to calculate the matrix of the modulation transfer functions (mtf) for each octave band (see equation 10), and then the speech transmission indices (STI), $L_{N,k}$ values were set using the levels obtained by the Algorithms 1 and 2. In this way, it was possible "to map" the speech privacy criteria for each source-receiver combination (Dickschen *et al.*, 2018).

IV. RESULTS

Tables I and II include the results of first step of clustering for both Algorithms. The candidate noise sources are sought here. For Algorithm 1, this is achieved by looking for the lowest AIC value in each octave band. Algorithm 2 is applied with the same number of clusters used in Algorithm 1 in order to make a fair comparison. The standard deviations of the Algorithm 1 are shown in brackets to have an overview of the clusters' density.

The intermediate values achieved in the second step are shown in Figure 3. The clustering evaluation metrics – AIC for Algorithm 1 and silhouettes for Algorithm 2 – assess the two unsupervised techniques for a number of clusters K from 1 to 10 for AIC and 2 from 10 for silhouettes. Even though the analyses were carried out for K up to 10, the most significant results are plotted up to $K = 6$ for a better visualization. $K = 2$ represents the best candidate model for both metrics. For Algorithm 1, since AIC has an asymptotical tendency, the elbow of the curves represents the reference for the best outcome. The 2 and 4 kHz octave bands reveal a slight change of slope between $K = 2$ and $K = 3$, but not significant. For Algorithm 2, silhouettes show the high coefficients for $K = 2$ in each octave band.

Then, the final outcomes produced by both Algorithms in the third step are shown in Table III. Now the candidate sound sources are labeled as either mechanical or human, on the basis of the above mentioned hypotheses (see Section III B). Bottalico and Astolfi measured vocal doses of elementary male and female teachers finding an uncertainty of the mean of the SPLs of about 4 dB (Bottalico and Astolfi, 2012). Olsen measured a standard deviation of the mean of speech in the range of 4-6 dB (Olsen, 1998). Iannace et al. measured the mechanical noise within an open-plan office in three operating conditions: two different speeds and the background noise with the HVAC system off. The standard deviations in the first two cases were about 1 dB, in the third was about 4 dB (Olsen, 1998). Leonard and Chilton reported the measured ambient noise levels of previous studies in open-plan offices. It is shown how the difference between minima and maxima SPLs span between 5 and and 11 dB (Peter and Anthony, 2019).

Concerning the results of the present work, for Algorithm 1 the distinction is made setting a standard deviation equal to 5 dB as threshold. If for a given sound source the standard deviation is smaller than 5 dB, then this source is classified as mechanical, otherwise it is associated to human activity. The standard deviations and the mixing proportions of each Gaussian curve are shown in brackets. Regarding the Algorithm 2, the identification is carried out analyzing the average distance of data points from the centroid within a cluster. These distances and the size of clusters, represented as a percentage of the whole data population, are shown in brackets. A short distance means a low spread of data points within the cluster, which is referable to a mechanical source. A large distance highlights a dynamic behaviour of the source, thus it is reasonable to associate this kind of source to human activity. The clusters percentage breakout indicates that a large amount of data belong to the mechanical source, on average 79% on the whole population. Consequently, the office under study can be considered as a quiet environment. On the bottom of Table III, the equivalent L_{eq}, and the percentile levels L_{90}, L_{50}, and L_{10} have been reported for comparison.

V. DISCUSSION

The office under study is located far from the city center in a quiet area, so indoor noises were expected to be the main components of the monitored soundscape (Acun and Yilmazer, 2018). In particular, they are the noise due to service equipments and office devices, and the human noise. The results of unsupervised analyses confirm this intuition: indeed the clustering evaluation finds $K = 2$ as the best model among candidates. Now, it has to be ascertained wether these numbers have a physical sense.

TABLE I. Results of the first step for Algorithm 1. Here, the candidate noise sources are sought. The results are obtained taking into account the lowest AIC possible. The standard deviations of the Gaussian curves are shown in brackets. All values are presented in dB for each octave band.

Frequency octave band (Hz)					
125	250	500	1000	2000	4000
Algorithm 1					
–	–	–	–	17.2 (0.4)	19.0 (0.4)
–	–	–	–	17.9 (0.3)	20.8 (0.3)
–	–	23.5 (1.2)	–	18.6 (0.4)	21.2 (0.2)
–	26.8 (1.4)	25.7 (1.1)	19.7 (1.0)	19.6 (0.6)	21.7 (0.2)
–	28.5 (1.0)	27.9 (1.0)	21.3 (0.8)	21.1 (0.9)	22.1 (0.3)
–	30.5 (1.0)	30.2 (1.1)	23.1 (1.1)	23.2 (1.2)	22.8 (0.7)
31.1 (1.9)	32.9 (1.4)	33.1 (1.7)	25.9 (1.9)	26.3 (1.7)	24.4 (1.6)
34.6 (3.3)	36.2 (2.6)	37.1 (2.9)	31.2 (3.9)	30.4 (3.0)	27.9 (3.2)
44.2 (6.9)	45.2 (6.8)	46.3 (7.9)	39.0 (8.3)	35.8 (7.1)	34.2 (6.0)

TABLE II. Results of the first step for Algorithm 2. Here, the candidate noise sources are sought. Algorithm 2 is applied with the same number of clusters set for Algorithm 1 in order to make a fair comparison. The average distances between each data point and the centroid within a cluster are shown in brackets. All values are presented in dB for each octave band.

Frequency octave band (Hz)					
125	250	500	1000	2000	4000
Algorithm 2					
–	–	–	–	17.8 (0.40)	20.6 (0.36)
–	–	–	–	19.9 (0.48)	21.8 (0.15)
–	–	23.9 (1.6)	–	22.6 (0.68)	23.4 (0.31)
–	27.0 (1.57)	27.8 (1.35)	20.6 (1.35)	25.7 (0.82)	25.7 (0.49)
–	30.6 (1.19)	32.0 (1.58)	24.3 (1.46)	28.9 (0.98)	28.3 (0.67)
–	34.6 (1.69).	36.5 (2.23)	29.0 (2.22)	32.5 (1.32)	31.4 (0.96)
31.3 (3.58)	39.8 (2.98)	42.4 (3.50)	34.7 (3.29)	36.8 (1.91)	35.2 (1.49)
37.7 (5.79)	46.8 (4.17)	49.5 (4.97)	42.0 (6.10)	42.2 (3.28)	40.0 (2.99)
49.6 (17.70)	54.0 (10.35)	58.0 (11.63)	52.1 (18.82)	49.8 (12.10)	47.7 (13.76)

TABLE III. Results of the third step for both Algorithms. The final outcomes associated to mechanical or human sources are shown. They are obtained running both Algorithms with $K = 2$, the optimal number of clusters found in the second step through the evaluation clustering (see Figure 3). For Algorithm 1, the standard deviations and the mixing proportions of the Gaussian curves of the Gaussian curves are shown, in brackets. For Algorithm 2, the average distances between each data point and the centroid within a cluster, and the size of each cluster expressed as percentage on the whole population, are shown, in brackets. Lastly, the equivalent and the 10th, 50th, and 90th percentile levels are shown for comparison. All values are presented in dB for each octave band.

Source type	Frequency octave band (Hz)					
	125	250	500	1000	2000	4000
Algorithm 1 – K=2						
Mech. (L_B)	32.5 (2.7 – 0.73)	30.0 (3.1 – 0.67)	28.1 (3.9 – 0.65)	22.2 (2.3 – 0.53)	18.6 (1.3 – 0.44)	21.6 (0.8 – 0.61)
Human (L_S)	41.6 (7.0 – 0.27)	41.3 (7.7 – 0.33)	40.7 (9.1 – 0.35)	32.5 (7.8 – 0.47)	28.0 (6.9 – 0.56)	27.5 (5.4 – 0.39)
Algorithm 2 – K=2						
Mech. (L_B)	32.7 (8.25 – 82%)	30.6 (11.75 – 79%)	28.6 (17.15 – 77%)	23.4 (10.16 – 75%)	20.3 (7.97 – 74%)	22.2 (3.01 – 84%)
Human (L_S)	45.5 (28.27 – 18%)	45.8 (31.65 – 21%)	45.8 (46.72 – 23%)	37.8 (40.45 – 25%)	33.7 (32.52 – 26%)	32.4 (21.11 – 16%)
L_{10}	42.9	45.4	45.5	37.7	34.0	30.0
L_{50}	33.3	31.7	30.3	24.6	21.3	22.1
L_{90}	29.5	26.7	23.7	19.9	17.5	20.7
L_{eq}	42.6	44.1	46.2	40.1	34.6	30.3

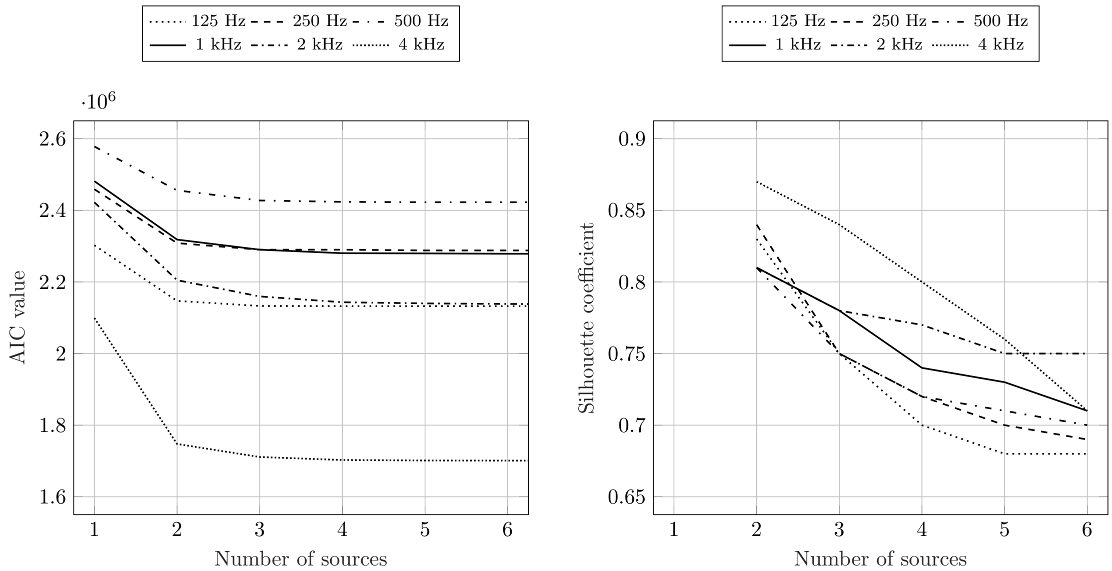


FIG. 3. Results of the second step for both Algorithms. AIC values on the top and silhouette evaluation on the bottom are shown for each octave band. The lower the AIC the better the model. The elbow of these curves represents the reference to evaluate the proper number of clusters to take into account in the analysis. Concerning the silhouettes, the higher the coefficient the better the model.

The two initial hypotheses used to identify the kind of source – i.e. mechanical or human – are similar but have different concepts and applications. The continuous and constant activity of the mechanical systems and devices is addressed with two different approaches: the statistical occurrences in Algorithm 1, and the number of bins in Algorithm 2.

The unsupervised algorithms are able to detect different densities within the SPLs collected by a sound level meter, regardless of its nature, mechanical or human. Finding the centre of gravity, like means or centroids, implies the capability of quantifying the sources in a dynamic context.

In the following, the results of both Algorithms will be compared from a statistical and spectral point of view, respectively.

A. Statistical remarks on the sound sources

After finding the clusters, hence the active sound sources, it is necessary to label them, i.e. to associate each cluster to an existing sound source. The statistical approach of Algorithm 1 has more features to investigate in order to find the metrics and describe the nature of the source. The metrical approach of Algorithm 2 needs to find similar metrics in order to compare the results. In this regard, a small standard deviation means that the associated sound source produces stable sound pressure levels continuously during time. It is reasonable to associate this kind of behaviour to a mechanical source. In contrast, a high standard deviation means a more accidental nature of the sound source, like human activity (Bottalico and Astolfi, 2012). Similarly, a cluster shaped by a mechanical source should have a high density of data points. This means a short average distance among the data points and their centroid within the cluster. A more random source should have more spread sound pressure levels, thus a lesser density and a larger average distance among data points and the centroid. These considerations are confirmed by the fact that standard deviations from Algorithm 1 and average distances of data points from Algorithm 2 have the same trend. Moreover, the mixing proportions of clusters is quite similar as well. In fact, the absolute values of the weight of each cluster take different values in the two Algorithms, but in both methods the mechanical source has the higher weight. Just one exception is present, in the 2 kHz band.

GMM can be considered as a generalization of KM for very small variances (MacKay, 2003). The higher the variances the larger the differences between the values achieved by Algorithm 1 and 2. Thus, this can be considered as a consequence of the heteroscedasticity of data.

The large population of the recorded data seems to give more robustness to Algorithm 1. Concerning this point, in Figure 4 the coefficients of variation, i.e. the ratio between the standard deviations and the mean values, are plotted for each octave band and for the two kind of sources, previously identified as mechanical and human. These coefficients show the dispersion of the data distribution for Algorithm 1. The trend is the same up to the 500 Hz octave band. Beyond this point, the gap between the curves of the mechanical and human coefficients of variation increases. The spread of the human activity noise increases up to the 2 kHz octave band: this source increases its dynamical behaviour in a range crucial for the human speech, where most formants occur. In the 4 kHz band there is a change of

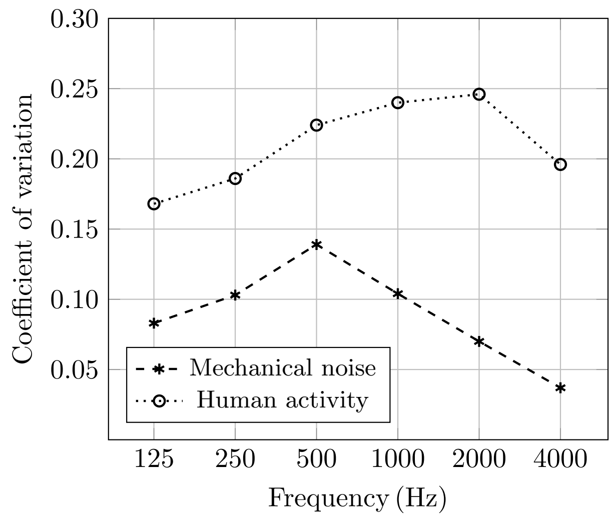


FIG. 4. Coefficient of variation of the two sources, the mechanical noise and the human activity, obtained by Algorithm 1. Results are plotted for each octave band.

tendency. So, this general trend supports the identification of the higher curve as belonging to the human activity.

B. Spectral remarks

The human nature of the higher means and centroids resulting from the unsupervised analysis can be confirmed by a spectral matching technique. Thus, the sound pressure levels of speech in the office under study, was calculated using the diffuse field hypotheses and the in situ measured values of reverberation time (Hodgson *et al.*, 2007). In fact, all the distances among workstations are greater than the critical distance of the room. A talking time of about the 20% of the whole monitoring time was considered. Because the background noise levels, measured within the office, remain almost below 45 dB, the Lombard effect is not triggered; this allows to use a constant value of speech power level (Peter and Anthony, 2019). The sound power level of normal speech has been set, according to the ISO 3382-3, as an averaged value between male and female speakers and for a normal voice effort. Taking into account recent findings, it is worth noting that the speech spectrum may change in noisy environments, especially on lower bands (Leembruggen *et al.*, 2016; Rindel *et al.*, 2012). In light of this, the 125 and 250 Hz octave bands of the ISO speech spectrum have been increased respectively of 6 and 3 dB. The SPLs calculated in this way were then compared with the measured values of human noise obtained with Algorithm 1, Algorithm 2, and L_{eq} (see Figure 5). The human activity is not continuous in each recorded frame; it represent just a percentage of time of the whole data population collected. In Figure 5 the dashed curve refers to the expected speech spectrum.

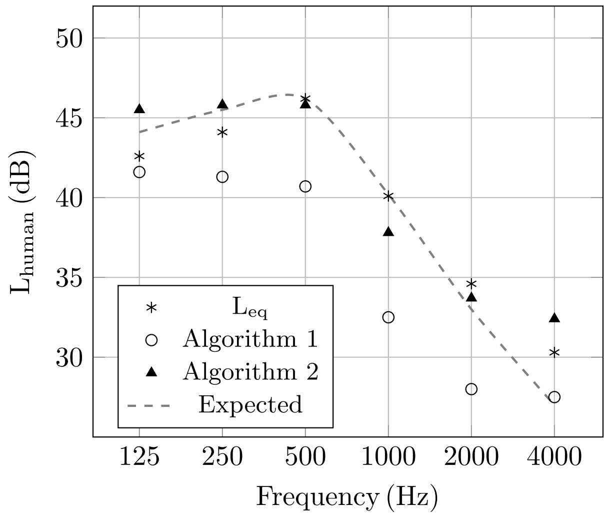


FIG. 5. Spectral matching between the calculated speech level in the room (under a diffuse sound field assumption; dashed line) and inferred values of the noise source identified as human using Algorithms 1 and 2 in the office under study. The dashed line is plotted assuming the speech running for the 20% of the monitoring time.

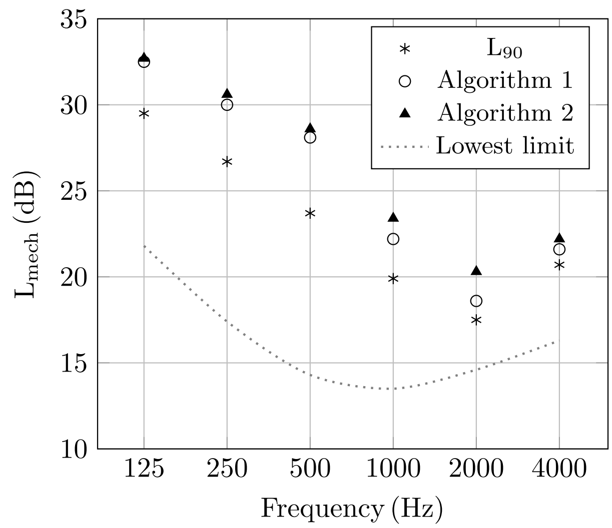


FIG. 6. Inferred values of the noise source identified as mechanical using Algorithms 1, Algorithm 2 compared with measured L_{90} in the office under study. The dotted curve shows the lowest detectable limit of the equipment used.

The qualitative analysis of the results shows how the detected human activity has a similar spectrum of the calculated speech in the office. The quantitative analysis shows differences among the methods. The Algorithm 1 gives significantly lower values compared to other methods and the expected curve. This gap among values could concern the type of classification of the clustering techniques. The GMM is defined as a soft clustering technique, i.e. each data point is assigned to each cluster with different probabilities, whereas the KM is an hard clustering technique, thus each data point can be assigned to one and only one cluster (Saxena *et al.*, 2017). Thus, the overlap zone between two Gaussian curves obtained by a mixture seems to lower the SPLs attributable to the source, corresponding to the mean values of the Gaussian curves. The hard clustering performed by the Algorithm 2, seems to process the classification similarly to the equivalent levels and the expected spectrum calculated instead. The quite flat tendencies on lower frequencies obtained via Algorithms 1 and 2 could be due to the masking effects of the mechanical noise, which is louder in this octave band, with respect to the few energy of the speech spectrum in these octave bands. Such effect has been already noticed by (D'Orazio *et al.*, 2020). Nevertheless, the flat tendency at the low frequencies falls within the uncertainties mentioned above.

Concerning Algorithm 1, this outcome can be led back to the low offset noticed in Figure 4. The separation in the lower band of 125 and 250 Hz seems to be more challenging, maybe due to the high energetic contribution of both sources, mechanical and human. The comparison between the mean values identified as human source and the diffuse speech levels characteristic of the office under study shows the same trend with differences up to 4 dB. This result confirms, with good approximation, the threshold chosen for the standard deviation. Another confirmation of the reliability of the Algorithms 1 and 2 is obtained looking at the clusters obtained by Algorithm 2. In fact, the average percentage over all the octave band of the human activity is about 21% of the monitoring time (see Table III). This can be assumed, in a first approximation, as the percentage of speech occurrence in the office during work hours. Thus, the trend of the Algorithm 2, shown in Figure 5, seems to be the more similar to the energetic model, since it is near the expected curve of the speech for the 20% of the whole monitoring time.

Moreover, in Fig. 6 the spectra of the mechanical noise obtained by Algorithms 1 and 2 and the 90th percentile levels are shown. The small differences between Algorithm 1 and Algorithm 2, as stated in the previous section, can be explained as a consequence of the small heteroscedasticity, and thus the low variances of the data, between the mechanical sources obtained via GMM and KM. Moreover, the mechanical noise measured via Algorithms 1 and 2 has greater values than the 90th percentile level usually used.

An unforeseen tendency is presented by the 4000 Hz octave band in both spectra, mechanical and speech. In fact, it is expected a strong decrease of these values for both sources. The dotted line represents the lowest limit detectable of the sound level meter used. Looking at this and considering the high quiet of the office, the growth of the levels in the 4000 Hz band of the mechanical spectrum, as well as the small decrease of this octave band in the speech spectrum, can be given to the intrinsic error of the instrument. This observation seems to be confirmed by the behaviour of the AIC and the silhouette coefficient of the 4000 Hz octave band shown in Figure 3. The large gap of the AIC value and the different tendency of the silhouette coefficient, compared to the other octave bands, suggests a different distribution of the SPLs within the database, hence imputable to the intrinsic noise of the measurement.

C. Background noise correction for STI evaluation

A simulation of the office was done to obtain the STI values without background noise, STI_{∞}, and then they were corrected with the background noise levels obtained with the two Algorithms and the percentile levels (table III). STI values, corrected with the contribution of mechanical and human noise separately, are shown in a source – receiver matrix (see Figure 7) where the gray scale reveals the variation of the parameter: from 0.5 (in black), which is the lower measured value, to 1 (in white), which is the ideal value of perfect intelligibility achieved when source and receiver are in the same workstation. The correction of STI_{∞} was made in two steps: at first, only the background noise due to the mechanical sources has been considered; then the contribution of the human activity was added . In figure 7 it is possible to see that there is a slight difference in the gray shade between the first two matrices in a row and only in the third matrix, when both the types of noise are considered, the shades are darker. For Algorithm 1, the variation of STI values is noticeable when only the mechanical contribution or the human contribution is considered. These results highlight that a more detailed analysis of the background noise allows to better evaluate the variation of STI values, avoiding to overestimate this important quantity.

These results suggest how to assess the effective privacy condition within the office, which is quite different than the privacy condition measured with the mechanical noise only, as currently required by technical standards. In fact, the effective privacy is significantly

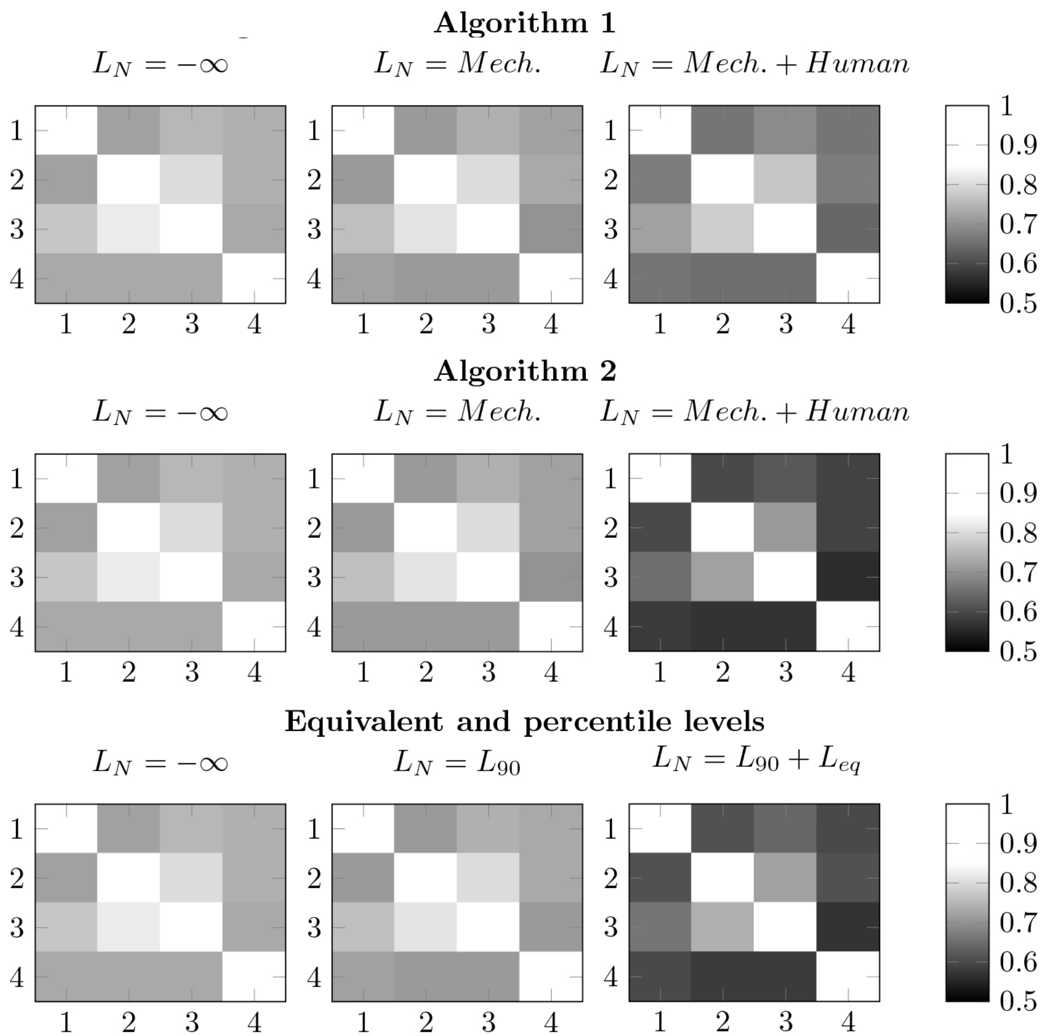


FIG. 7. Matrices of the STI values among the workstations in the office under study. The source has been set at the "normal" speech level. On each row, going from the left to right, the STI is presented first without background nose (indicated as $L_N = -\infty$), then corrected with the background noise levels obtained through the unsupervised analysis. First adding the mechanical contribution only (indicated as $L_N = L_{mech}$), then summing up the human contribution as well (indicated as $L_N = L_{mech} + L_{human}$). The sidebar on the right represents the legend of the STI values. On the axis of the matrix are reported the source-receiver positions $(1-4)$ corresponding to the three workstations and the meeting table (see Figure 2).

affected by the social context (Rasmussen and Carrascal García, 2019). Where the active noise masking is used, the speech privacy can be quite constant over the working areas. In other cases, such as the one under study, the privacy fluctuates dynamically in time. The results obtained with the procedure presented in this work allow to assess different scenarios, thus broadening the characterization of privacy criteria of ISO 3382-3. Further analyses could be done with these unsupervised Algorithms with longer monitoring times, in order to investigate the existence of more significant correlations with percentile levels (Renz *et al.*, 2019).

VI. CONCLUSIONS

Workers' comfort and productivity inside offices is influenced by the background noise, which is due to different contributions (mechanical equipment, outdoor traffic, human activity). Therefore it is highly desirable to be able to separate these contributions from the temporal history provided by a simple monitoring systems. This would allow to control HVAC noise during working hours or to dynamically optimize the speech privacy between workers. Therefore, unsupervised algorithms capable to perform this task in an automated manner are required. However, even if some procedures were proposed in previous research, the instance seems to be still open. In the present work, two unsupervised methods capable to separate and identify different noise sources from the same recording are described in details. Both are based on clustering algorithms (GMM and K-Means) and further refined by a clustering evaluation. The third step of each Algorithm may be adapted on the context under study, on the basis of statistical and temporal preliminary observations.

The two Algorithms have been checked on a dataset of short-time (100 ms) equivalent levels, coming from the recording of a whole working day in a small office. Two noise sources, later identified as mechanical and human noise, were extracted from the dataset. It has been shown how the candidates noise sources - extracted in the first step - were easily reduced by the clustering evaluation. The Akaike information criterion and the silhouette criterion where applied in each octave band, returning comparable results. It was noted that Algorithm 1 is more sensitive to the statistical characteristics of the noise sources, while Algorithm 2 is more sensitive to the temporal behavior. As a consequence, in the third step, the proposed Algorithms return slightly different results. Indeed the two kind of noise sources vary less (the mechanical one) or more (the human one) in time, so that the homoscedasticity is not reached. Therefore, there are uncertainties at low frequencies, where the speech noise energy is lower than the mechanical noise one. Increasing the frequency of the input signal, the two sound sources seem to be identified by the statistical-based approach (Algorithm 1) better than by the metric-based approaches (Algorithm 2). Nevertheless, Algorithm 2 return information on the temporal behavior that are useful to optimize Algorithm 1, so both of them seems to be needed for in a depth analysis.

With respect to previous researches, the unsupervised Algorithms presented here are quite robust and, after a preliminary set, they could be implemented, e.g., in continuous monitoring systems.

ACKNOWLEDGMENTS

The authors would like to thank Elena Rossi for the help during the preliminary research, Giulia Fratoni for having set the numerical model, Laura Anderlucci for the useful hints on the statistical analysis and Michele Ducceschi for the kind revision.

Acun, V., and Yilmazer, S. (**2018**). "A grounded theory approach to investigate the perceived soundscape of open-plan offices," Applied Acoustics **131**, 28–37.

Akaike, H. (**1974**). "A new look at the statistical model identification," IEEE transactions on automatic control **19**(6), 716–723.

Alimohammadi, I., and Ebrahimi, H. (**2017**). "Comparison between effects of low and high frequency noise on mental performance," Applied Acoustics **126**, 131–135.

Bianco, M. J., Gerstoft, P., Traer, J., Ozanich, E., Roch, M. A., Gannot, S., and Deledalle, C.-A. (**2019**). "Machine learning in acoustics: Theory and applications," The Journal of the Acoustical Society of America **146**(5), 3590–3628.

Bottalico, P., and Astolfi, A. (**2012**). "Investigations into vocal doses and parameters pertaining to primary school teachers in classrooms," The Journal of the Acoustical Society of America **131**(4), 2817–2827.

Braat-Eggen, E., vd Poll, M. K., Hornikx, M., and Kohlrausch, A. (**2019**). "Auditory distraction in open-plan study environments: Effects of background speech and reverberation time on a collaboration task," Applied Acoustics **154**, 148–160.

Dehlbæk, T. S., Brunskog, J., Petersen, C. M., and Marie, P. (**2016**). "The effect of human activity noise on the acoustic quality in open plan offices," in *INTER-NOISE and NOISE-CON Congress and Conference Proceedings*, Institute of Noise Control Engineering, Vol. 253, pp. 4117–4126.

Dempster, A. P., Laird, N. M., and Rubin, D. B. (**1977**). "Maximum likelihood from incomplete data via the em algorithm," Journal of the Royal Statistical Society: Series B (Methodological) **39**(1), 1–22.

Di Blasio, S., Shtrepi, L., Puglisi, G. E., and Astolfi, A. (**2019**). "A cross-sectional survey on the impact of irrelevant speech noise on annoyance, mental health and well-being, performance and occupants' behavior in shared and open-plan offices," International journal of environmental research and public health **16**(2), 280.

Dickschen, A., Bertazzoni, R., and Liebl, A. (**2018**). "Evaluating room acoustic quality in open-plan offices by adding the distribution of possible source and receiver positions to the simulation," Proceedings Euronoise 2018 27–31.

D'Orazio, D., De Salvio, D., Anderlucci, L., and Garai, M. (**2020**). "Measuring the speech level and the student activity in lecture halls: Visual-vs blind-segmentation methods," Applied Acoustics **169**, 107448.

D'Orazio, D., Rossi, E., and Garai, M. (**2018**). "Comparison of different in situ measurements techniques of intelligibility in an open-plan office," Building Acoustics **25**(2), 111–122.

Ebissou, A., Parizet, E., and Chevret, P. (**2015**). "Use of the speech transmission index for the assessment of sound annoyance in open-plan offices," Applied Acoustics **88**, 90–95.

Ellermeier, W., Eigenstetter, M., and Zimmer, K. (**2001**). “Psychoacoustic correlates of individual noise sensitivity,” The Journal of the Acoustical Society of America **109**(4), 1464–1473.

Ellermeier, W., and Zimmer, K. (**2014**). “The psychoacoustics of the irrelevant sound effect,” Acoustical Science and Technology **35**(1), 10–16.

Gabbay, D. M., Thagard, P., Woods, J., Bandyopadhyay, P. S., and Forster, M. R. (**2011**). *Philosophy of Statistics*, **7** (Elsevier).

Harvie-Clark, J., Larrieu, F., and Opsanger, C. (**2019**). “Iso 3382-3: Necessary but not sufficient. a new approach to acoustic design for activity-based-working offices,” Proceedings of the 23rd International Congress on Acoustics .

Hodgson, M., Rempel, R., and Kennedy, S. (**1999**). “Measurement and prediction of typical speech and background-noise levels in university classrooms during lectures,” The Journal of the Acoustical Society of America **105**(1), 226–233.

Hodgson, M., Steininger, G., and Razavi, Z. (**2007**). “Measurement and prediction of speech and noise levels and the lombard effect in eating establishments,” The Journal of the Acoustical Society of America **121**(4), 2023–2033.

Hongisto, V. (**2005**). “A model predicting the effect of speech of varying intelligibility on work performance.,” Indoor air **15**(6), 458–468.

Iannace, G., Ciaburro, G., and Trematerra, A. (**2018**). “Heating, ventilation, and air conditioning (hvac) noise detection in open-plan offices using recursive partitioning,” Buildings **8**(12), 169.

IEC 60268 (**2020**). “Sound System Equipment - Part 16: Objective rating of speech intelligibility by speech transmission index” (International Electrotechnical Commission, Geneva, Switzerland).

ISO 22955 (**2020**). “Acoustics - Acoustic quality of open office spaces” (International Organization for Standardization, Geneva, Switzerland).

ISO 3382 (**2012**). “Acoustics - Measurement of room acoustic parameters — part 3: Open-plan offices” (International Organization for Standardization, Geneva, Switzerland).

Jauhiainen, S., and Kärkkäinen, T. (**2017**). “A simple cluster validation index with maximal coverage,” in *European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning*, ESANN.

Jones, D. (**1999**). “The cognitive psychology of auditory distraction: The 1997 bps broadbent lecture,” British Journal of Psychology **90**(2), 167–187.

Koskela, H., Maula, H., Haapakangas, A., Moberg, V., and Hongisto, V. (**2014**). “Effect of low ventilation rate on office work performance and perception of air quality–a laboratory study,” Proceedings of Indoor Air 673–675.

Leembruggen, G., Verhave, J. *et al.* (**2016**). “The effect on sti results of changes to the male test-signal spectrum,” Proc IOA **38**, 78–87.

Lloyd, S. (**1982**). “Least squares quantization in pcm,” IEEE transactions on information theory **28**(2), 129–137.

MacKay, D. J. (**2003**). *Information theory, inference and learning algorithms* (Cambridge university press).

McLachlan, G.J. and Peel, D. (**2000**). *Finite Mixture Models* (Wiley Series in Probability and Statistics, New York).

Olsen, W. O. (**1998**). "Average speech levels and spectra in various speaking/listening conditions," American Journal of Audiology **7**, 21–25.

Perrin Jegen, N., and Chevret, P. (**2017**). "Effect of noise on comfort in open-plan offices: application of an assessment questionnaire," Ergonomics **60**(1), 6–17.

Peter, L., and Anthony, C. (**2019**). "The lombard effect in open plan offices," in *Proceedings of the Institute of Acoustics*, Institute of Acoustics, Vol. 41, pp. 216–226.

Rasmussen, B., and Carrascal García, T. (**2019**). "Acoustic regulations for offices-comparison between selected countries in europe," in *INTER-NOISE and NOISE-CON Congress and Conference Proceedings*, Institute of Noise Control Engineering, Vol. 259, pp. 8141–8150.

Renz, T., Leistner, P., and Liebl, A. (**2018**a). "Auditory distraction by speech: Can a babble masker restore working memory performance and subjective perception to baseline?," Applied Acoustics **137**, 151–160.

Renz, T., Leistner, P., and Liebl, A. (**2018**b). "Auditory distraction by speech: Sound masking with speech-shaped stationary noise outperforms- 5 db per octave shaped noise," The Journal of the Acoustical Society of America **143**(3), EL212–EL217.

Renz, T., Leistner, P., and Liebl, A. (**2019**). "Use of energy-equivalent sound pressure levels and percentile level differences to assess the impact of speech on cognitive performance and annoyance perception," Applied Acoustics **153**, 71–77.

Rindel, J. H. (**2018**). "Open plan office acoustics–a multidimensional optimization problem," Proceedings of DAGA2018, Munich, Deutsche Gesellschaft für Akustik .

Rindel, J. H., Christensen, C. L., and Gade, A. C. (**2012**). "Dynamic sound source for simulating the lombard effect in room acoustic modeling software," in *INTER-NOISE and NOISE-CON Congress and Conference Proceedings*, Institute of Noise Control Engineering, Vol. 2012, pp. 954–966.

Rodríguez, C. C. (**2005**). "The abc of model selection: Aic, bic and the new cic," in *AIP Conference Proceedings*, American Institute of Physics, Vol. 803, pp. 80–87.

Rousseeuw, P. J. (**1987**). "Silhouettes: a graphical aid to the interpretation and validation of cluster analysis," Journal of computational and applied mathematics **20**, 53–65.

Saxena, A., Prasad, M., Gupta, A., Bharill, N., Patel, O. P., Tiwari, A., Er, M. J., Ding, W., and Lin, C.-T. (**2017**). "A review of clustering techniques and developments," Neurocomputing **267**, 664–681.

Schlittmeier, S. J., and Liebl, A. (**2015**). "The effects of intelligible irrelevant background speech in offices–cognitive disturbance, annoyance, and solutions," Facilities .

Schwarz, G. (**1978**). "Estimating the dimension of a model," Annals of statistics **6**(2), 461–464.

Vellenga, S., Bouwhuis, T., and Höngens, T. (**2017**). "Proposed method for measuring 'liveliness' in open plan offices," in *Proceedings of the 24th International Congress on Sound and Vibration, London, UK*, pp. 23–27.

Vorländer, M. (**2020**). *Auralization* (Springer).

Wang, L. M., Brill, L. C., Nord, J., and Bovaird, J. (**2020**). "Higher speech levels in k-12 classrooms correlate with lower math achievement scores," The Journal of the Acoustical Society of America **148**(4), 2631–2632.

Yadav, M., and Cabrera, D. (**2019**). "Two simultaneous talkers distract more than one in simulated multi-talker environments, regardless of overall sound levels typical of open-plan offices," Applied Acoustics **148**, 46–54.

Yadav, M., Kim, J., Cabrera, D., and De Dear, R. (**2017**). "Auditory distraction in open-plan office environments: The effect of multi-talker acoustics," Applied Acoustics **126**, 68–80.