Tagging noise sources in offices through Machine-Learning techniques

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

[1]DIN University of Bologna, Italy

*{dario.dorazio@unibo.it}

Abstract
Measuring and monitoring background noise levels in active offices is a crucial issue to assess comfort conditions and enhance the technical standards. Machine learning-based algorithms have been proposed in previous work to identify and measure noise sources during working days in a test office. Two unsupervised clustering techniques, i.e. Gaussian Mixture Model and K-means clustering, gathered the recorded sound pressure levels finding patterns. After the clustering, preliminary labelling of the sources (i.e., human, mechanical, traffic, and so on) was made employing statistical and metric assumptions. In the present work, the method has been tested in a wider scenario: three offices were recorded during two working days. Thus, the ML-based algorithms were run finding contributions of the noise sources within the offices. The discrimination of noise sources was carried out by improved statistical and metric properties, which were briefly discussed. Results encourage testing ML algorithms in long-period monitoring.

Keywords: open-plan offices, k-means, gaussian mixture, clustering, acoustic monitoring.

1 Introduction

One of the main acoustic challenges of the recent ISO 22955:2021 is to limit the worker's exposure to ambient noise. A-weighted, equivalent continuous sound level $L_{Aeq,T}$ was proposed by ISO 22955:2021 to measure the workstation noise level. Target values were set for different types of activity (see table 1).

Many sound sources can contribute to workstation noise: human noise, mechanical noise, traffic noise from outdoor, ... In the field of open-plan office acoustics, the discrimination between noise sources is usually made with percentile levels. [1, 2].

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 supervision by the operator. For these reasons, unsupervised algorithms can represent a useful tool for accurate monitoring without the need for human supervision.

This work aims to identify sound sources via an unsupervised statistical analysis of long-term monitoring [3]. Data population obtained from the recording done in three active offices were processed with algorithms used in unsupervised learning to find patterns and create clusters. Then, sound sources were tagged by associating to each cluster - "human" or "mechanical" in these preliminary cases. Algorithm results were compared with equivalent continuous and percentile sound levels.

2 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.

Table 1: Workstation noise level $L_{Aeq,T}$ assumed for different types of activity, according to ISO 22955:2021.

Activities in its own space

Activity	Target values (dBA)
Activity mainly focusing on outside of the room communication	55
Activity mainly based on collaboration between people at the nearest workstation	52
Activity mainly based on a small amount of collaborative work	48
Activity can involve receiving public	55

Activities within the same space

Receiver space type	Target values (dBA)
Informal meetings (open plan)	48
Outside of the room communication (phone)	48
Collaborative	45
Noncollaborative	42
Focussed phone	42
Focussed individual work	40

3. Final clustering analysis and association of each cluster to a noise contribute based on 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 each Gaussian curve (see eqs. 1 and 2). In order to achieve meaningful results, the 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 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. 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. 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 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 provide a metric to evaluate the clustering goodness. Thus, the clustering validity is rated finding the highest silhouette coefficient SC (equation 9) among

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

candidate models, which means for a 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 with 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.

2.1. Clustering algorithms

Clustering algorithms allow the identification of 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 that decomposes the original model data into a sum of gaussian curves. Assuming a set of observations $x_1, \dots, x_n$ (e.g. the short-time equivalent levels recorded), the Gaussian probability density function $f(x_i)$ of these observations – in 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* [4], non-negative quantities that sum to one; that is,

$$0 \le \pi_k \le 1 \quad (k = 1, \dots, K)$$

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

and

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

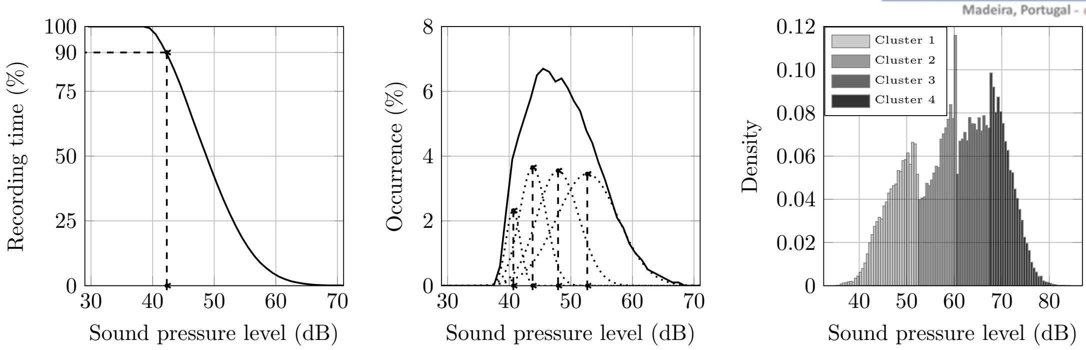
The likelihood function for a mixture model with K univariate Normal components is:

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

The equality in 1 is usually realized by the maximum likelihood optimization algorithm, e.g. the Expectation-Maximization (EM) [5]. In the context of background noise in open-plan offices, Dehlbaek et al. [6] 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 conditions.

While GMM is based on the statistical properties of the data population, KM optimizes a metric distance of each single data point 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:

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



(a) Percentile levels (b) Gaussian Mixture Model (c) K-means clustering

Figure 1: The three unsupervised methods used in this work. In figure (a) the continuous line represents 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 into 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 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 [7]. 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 [8].

2.2. 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 [9]. The clustering validation allows to assess the best model among candidates through specific metrics. In this work, the Akaike Information Criterion (AIC) [10] and the silhouette method [11] 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.

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 $s(i)$ express 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 datasets [12].

3 Method

In order to evaluate the performance of the two methods, the following scenarios were monitored:

Office A A mid-sized open plan: 8-12 workers varying during the day; Activity mainly focusing on outside of the room communication

Office B A mid-sized open plan: 10-12 workers varying during the day; Activity mainly based on a small amount of collaborative work

Office C A small-sized open plan: 0-2 workers varying during the day.

All open-plan offices were treated with COVID19 screen (120 cm height) without further acoustic improvements. The acoustic quality of all the spaces can be assumed as poor.

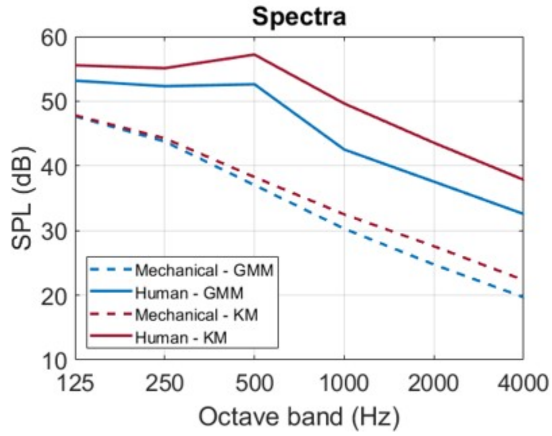
Sound pressure levels monitoring was carried out throughout two working days, so to allow to record enough data. A statistical data population was obtained by recording short-time equivalent levels, 100 ms integration time, octave-band filtered (125 Hz - 4000 Hz), for an amount of time long enough to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

The recordings were post-processed via two algorithms proposed in previous work [3]. The latter exploit two unsupervised clustering techniques commonly used in machine learning [13]: Gaussian Mixture Model (GMM) and K-means clustering (KM). First of all, the clustering validation step allows to set the best number of clusters available in the recorded population data. Each cluster represents a sound source, this is the assumption underlying this step. The metrics for the validation are the Akaike Informational Criterion (AIC) for GMM and the silhouette coefficient (SC) for KM [10, 11]. The corresponding elbow of the AIC curve and the greatest SC value are respectively the best number of clusters for each algorithm.

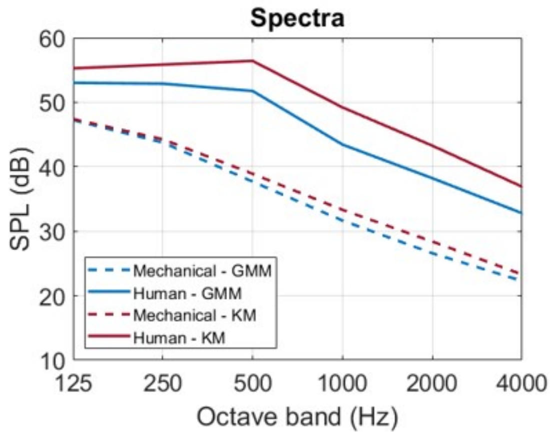
After the number of sound sources is discovered, their respective sound pressure levels (SPLs) are achieved via GMM and KM. GMM splits the data basing on probability assumptions, describing each cluster with a Gaussian curve. KM shapes the clusters basing on the minimization of the distance within the data [14]. In

Table 2: Results of the clustering carried out over long-term monitoring of the three offices. The office, the correspondent measurement day, the algorithm and the kind of source are shown. Measured SPLs are shown for each octave band from 125 to 4000 Hz, besides the A-weighted values. Moreover, the correspondent A-weighted continuous-equivalent level measured through the sound level meter is shown.

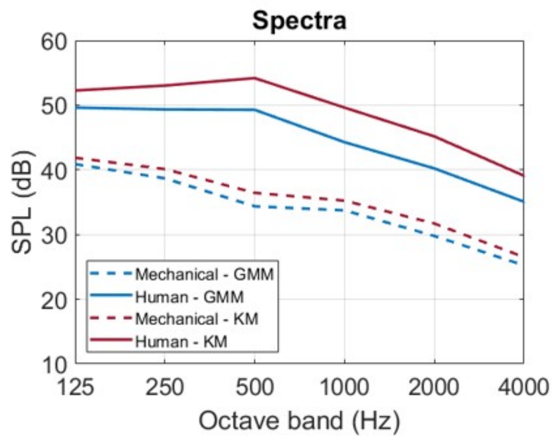
Office	Day	Algorithm	Source	Frequency octave band						A-weighted	$L_{Aeq,T}$
				125 Hz	250 Hz	500 Hz	1 kHz	2 kHz	4 kHz		
A	1	GMM	Mechanical	47.7 (1.7)	43.8 (1.4)	37.0 (1.2)	30.3 (1.3)	24.8 (2.3)	19.8 (3.7)	39.8 (1.0)	56.5
			Human	53.2 (4.3)	52.3 (5.3)	52.6 (8.5)	42.5 (8.6)	37.6 (8.5)	32.6 (8.4)	51.9 (7.7)	
		KM	Mechanical	47.8 (3.2)	44.3 (3.7)	38.3 (9.0)	32.5 (12.1)	27.6 (17.5)	22.3 (21.6)	41.3 (8.3)	
			Human	55.5 (9.5)	55.1 (15.6)	57.2 (34.3)	49.6 (41.2)	43.6 (40.8)	37.8 (36.0)	57.0 (29.6)	
A	2	GMM	Mechanical	47.2 (1.7)	43.8 (1.7)	37.7 (1.5)	31.7 (1.8)	26.6 (2.5)	22.3 (3.6)	40.2 (1.4)	55.4
			Human	52.9 (4.3)	53.0 (5.6)	51.8 (8.0)	43.4 (8.0)	38.2 (7.7)	32.8 (7.4)	51.8 (7.1)	
		KM	Mechanical	47.4 (3.2)	44.3 (4.6)	38.9 (8.9)	33.3 (11.1)	28.4 (13.9)	23.3 (14.9)	41.5 (7.8)	
			Human	55.3 (9.5)	55.8 (17.0)	56.4 (31.5)	49.2 (36.5)	43.3 (32.2)	36.9 (28.7)	56.3 (26.2)	
B	1	GMM	Mechanical	40.9 (2.2)	38.7 (2.4)	34.4 (2.2)	33.7 (2.1)	29.8 (2.3)	25.2 (3.1)	38.7 (2.1)	55.0
			Human	49.6 (5.1)	49.3 (6.4)	49.3 (8.3)	44.3 (7.4)	40.2 (7.0)	35.1 (6.8)	50.6 (7.1)	
		KM	Mechanical	41.9 (7.3)	40.1 (10.6)	36.4 (15.3)	35.2 (10.5)	31.7 (12.2)	26.5 (12.7)	40.7 (12.2)	
			Human	52.3 (13.3)	53.0 (20.6)	54.1 (33.1)	49.6 (32.2)	45.2 (27.1)	39.1 (25.0)	55.0 (25.7)	
B	2	GMM	Mechanical	41.7 (1.9)	38.2 (1.9)	33.3 (1.7)	32.7 (1.6)	28.7 (1.9)	23.8 (2.6)	37.9 (1.7)	54.5
			Human	49.1 (5.0)	48.2 (6.5)	47.9 (8.8)	43.2 (7.7)	39.3 (7.4)	33.6 (6.8)	49.5 (7.5)	
		KM	Mechanical	42.3 (5.1)	39.4 (8.2)	35.0 (1.2)	34.0 (8.9)	30.5 (10.5)	25.2 (11.3)	39.5 (9.9)	
			Human	51.8 (12.7)	52.3 (23.0)	53.6 (36.8)	49.3 (36.8)	44.9 (32.3)	37.9 (26.4)	54.6 (29.9)	
C	1	GMM	Mechanical	35.4 (3.2)	35.7 (3.2)	33.8 (4.8)	29.0 (4.2)	22.4 (3.7)	14.2 (1.1)	34.7 (3.5)	53.2
			Human	46.1 (7.4)	46.2 (7.5)	45.4 (8.8)	39.6 (8.1)	35.9 (8.3)	29.5 (8.5)	46.3 (7.9)	
		KM	Mechanical	36.2 (12.9)	36.5 (12.8)	34.2 (20.5)	30.3 (19.1)	25.9 (23.9)	20.6 (22.1)	36.6 (18.2)	
			Human	50.3 (27.9)	50.5 (29.6)	49.8 (40.8)	40.5 (32.0)	41.2 (32.2)	35.7 (34.2)	51.0 (32.0)	
C	2	GMM	Mechanical	34.7 (2.4)	33.5 (2.2)	31.4 (3.8)	27.7 (4.3)	22.0 (4.5)	16.9 (3.0)	33.5 (3.5)	51.7
			Human	43.8 (7.1)	44.1 (7.2)	45.5 (8.8)	38.6 (8.4)	34.7 (8.7)	29.6 (8.2)	45.7 (8.1)	
		KM	Mechanical	35.3 (8.8)	34.5 (9.2)	32.3 (18.1)	28.2 (17.7)	23.6 (23.4)	19.2 (18.2)	34.7 (17.0)	
			Human	48.3 (23.8)	48.8 (27.6)	50.0 (40.9)	43.0 (37.6)	39.7 (38.8)	34.8 (36.5)	50.4 (34.8)	



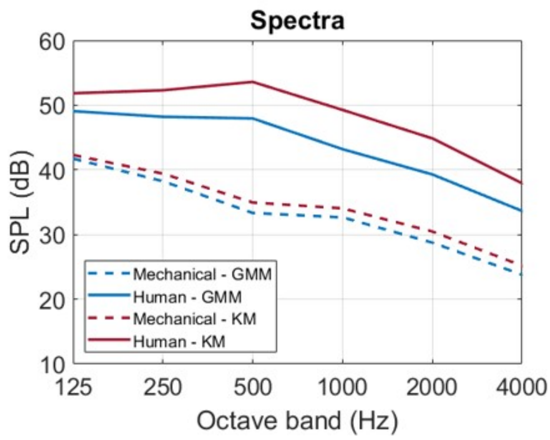
(a) **Office A** - Day 1



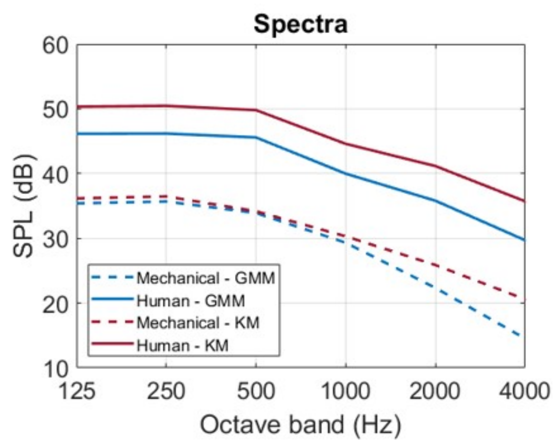
(b) **Office A** - Day 2



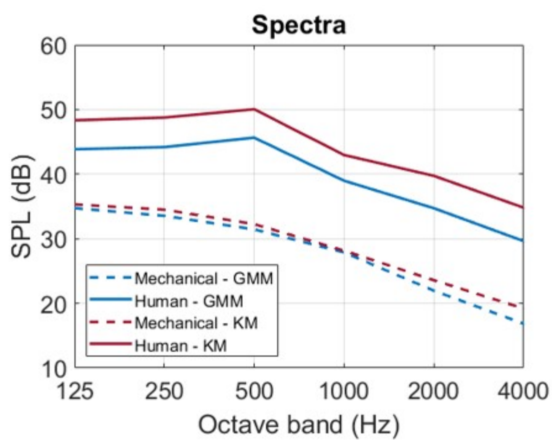
(c) **Office B** - Day 1



(d) **Office B** - Day 2



(e) **Office C** - Day 1



(f) **Office C** - Day 2

Figure 2: Spectral analysis: results of the spectra recorded in the three offices for day 1 (on the left) and day 2 (on the right).

this method, the means of the Gaussian curves and the centroids, i.e. the center of gravity of each cluster, are assumed as the SPLs produced by the sound sources.

The qualitative analysis concerns the detection of the kind of source, either *mechanical* or *human*. In the previously cited work [3], the identification of the kind of sound source has been made through the standard deviation (SD) of the Gaussian curves of each cluster obtained via GMM. With an SD lower than 5 dB, the source is labelled as mechanical, otherwise is considered due to human activity. Concerning the KM, a comparable metric of the s.d. was found in the average intra-cluster distance (AICD) among data points, i.e. the average distance of data points belonging to the same cluster.

4 Results

The outcomes concern the whole recorded working days of each office. The best number of clusters within the recorded data is $k = 2$ for offices A, B, and C; thus, the expected soundscape made up by two sound sources, as state above in the method section, is confirmed.

Once the number of clusters is set, the second step runs the algorithms, i.e. GMM and KM, and finds the SPLs of each sound source. Table 2 shows the results of GMM and KM of each office and day for the octave band from 125 up to 4000 Hz and the A-weighted equivalent level. Here, values are labelled as "mechanical" or "human" according to the threshold of 5 dB of the correspondent Gaussian curve. Brackets contain SD and AICD respectively for GMM and KM.

Furthermore, a preliminary method to assess the reliability of the algorithms is investigating the spectra obtained. Figure 2 shows all the spectra obtained in each office for each day and each algorithm in the octave bands from 125 to 4000 Hz. Blue and red lines indicate respectively the outcomes obtained by GMM and KM. Dashed and solid represent respectively mechanical and human sources.

Spectral behaviour and sound levels of identified sound sources are consistent with the physical scenarios in the measured offices. Indeed, the level of "human" source is influenced by the number of people within the office and its spectrum fits with speech power level, according to ISO 3382-3. Moreover, office equipment and HVAC can show a noise similar to ones measured through "mechanical" source. Furthermore, these methods can be used to measure the HVAC noise in dynamic scenarios [15].

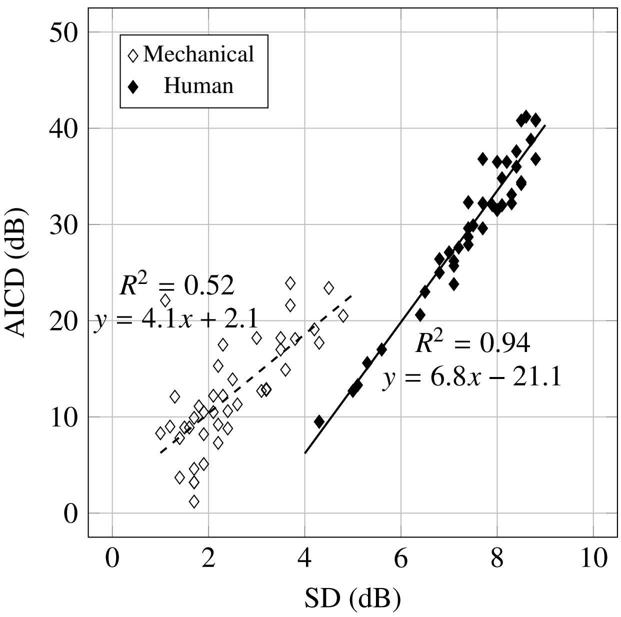


Figure 3: Relationship between the standard deviation SD and the average intra-cluster distance AICD. White and black diamonds represent respectively the mechanical and human noise detected during the monitoring.

The preliminary analysis of these three active offices seems to confirm threshold values used by algorithms to identify and label sources. Previous works about monitoring systems or human activities have shown an SD threshold of about 4-6 dB to split the two kinds of sources [16, 17]. The method used in the present paper sets the threshold to 5 dB. Values of SD smaller than 5 dB are deemed to belong to a *mechanical* sound source, otherwise to *human* ones. At the same time, it is important to find a similar metric for KM. Thus, a parameter that can describe the size of the cluster is the average distance of the data point within the same cluster, i.e. the intra-cluster distance AICD. Then, the correlation between SD and AICD helps to find the correspondent threshold by AICD. Figure 3 plots the measured SD and AICD for both mechanical and human sources. Mechanical sources appear more scattered, maybe because of the several devices within the offices; whereas the correlation concerning the human sources has an R^2 equal to 0.94. Except for the SD value measured in the 125 Hz octave band in office A during day 1, the SD equal to 5 dB seems to be confirmed as the threshold to split the two kinds of sources.

5 Conclusions

Two ML-based methods to measure more than one sound source in open-plan offices at the same time are presented. In the case under study, three open-plan offices were used to check both unsupervised methods. Sound sources were tagged, respectively, as 'mechanical' and 'human'. Spectral behaviour and sound levels of identified sound sources are consistent with the physical scenarios in the measured offices (number of people, office equipment, etc...). Results can usefully contribute to measuring the workstation noise level, as introduced by recent ISO 22955:2021.

References

[1] Sara Vellenga, Tom Bouwhuis, and Theodoor Höngens. Proposed method for measuring 'liveliness' in open plan offices. In *Proceedings of the 24th International Congress on Sound and Vibration, London, UK*, pages 23–27, 2017.

[2] Tobias Renz, Philip Leistner, and Andreas Liebl. 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, 2019.

[3] Domenico De Salvio, Dario D'Orazio, and Massimo Garai. Unsupervised analysis of background noise sources in active offices. *The Journal of the Acoustical Society of America*, 149(6):4049–4060, 2021.

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

[5] Arthur P Dempster, Nan M Laird, and Donald B Rubin. Maximum likelihood from incomplete data via the em algorithm. *Journal of the Royal Statistical Society: Series B (Methodological)*, 39(1):1–22, 1977.

[6] Tania Stenholt Dehlbæk, Jonas Brunskog, Claus Moller Petersen, and Pierre Marie. The effect of human activity noise on the acoustic quality in open plan offices. In *INTER-NOISE and NOISE-CON Congress and Conference Proceedings*, volume 253, pages 4117–4126. Institute of Noise Control Engineering, 2016.

[7] Stuart Lloyd. Least squares quantization in pcm. *IEEE transactions on information theory*, 28(2):129–137, 1982.

[8] Lily M Wang, Laura C Brill, Jayden Nord, and James Bovaird. 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, 2020.

[9] Gideon Schwarz. Estimating the dimension of a model. *Annals of statistics*, 6(2):461–464, 1978.

[10] Hirotugu Akaike. A new look at the statistical model identification. *IEEE transactions on automatic control*, 19(6):716–723, 1974.

[11] Peter J Rousseeuw. Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. *Journal of computational and applied mathematics*, 20:53–65, 1987.

[12] Susanne Jauhiainen and Tommi Kärkkäinen. A simple cluster validation index with maximal coverage. In *European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning*. ESANN, 2017.

[13] Michael J Bianco, Peter Gerstoft, James Traer, Emma Ozanich, Marie A Roch, Sharon Gannot, and Charles-Alban Deledalle. Machine learning in acoustics: Theory and applications. *The Journal of the Acoustical Society of America*, 146(5):3590–3628, 2019.

[14] Christopher M. Bishop. *Pattern Recognition and Machine Learning*. Springer, 2006.

[15] E Rossi, D De Salvio, D D'Orazio, and M Garai. Measuring and identifying background noises in offices during work hours. In *IOP Conference Series: Materials Science and Engineering*, volume 609, page 042005. IOP Publishing, 2019.

[16] Leonard Peter and Chilton Anthony. The lombard effect in open plan offices. In *Proceedings of the Institute of Acoustics*, volume 41, pages 216–226. Institute of Acoustics, 2019.

[17] Wayne O Olsen. Average speech levels and spectra in various speaking/listening conditions. *American Journal of Audiology*, 7:21–25, 1998.