Probabilistic Modelling of the Temporal Variability of Urban Sound Levels

Arnaud Can1), Pierre Aumond1), Bert De Coensel2), Carlos Ribeiro3), Dick Botteldooren2), Catherine Lavandier4)

1) LUNAM Université (l’Université Nantes-Angers-Le Mans), Ifsttar (Institut français des sciences et technologies des transports, de l’aménagement et des réseaux), AME-LAE (département Aménagement-Mobilité-Environnement, Laboratoire d’Acoustique Environnementale), 44341 Bouguenais, France. arnaudcan@ifsttar.fr
2) WAVES Group, Department of Information Technology, Ghent University, Technologiepark-Zwijnaarde 15, 9052 Gent, Belgium
3) Bruitparif, Axe Pleyel 4 - B104, 32 boulevard Ornano, 93200 Saint-Denis, France
4) ETIS, UMR 8051, Université Paris Seine, Université Cergy-Pontoise, ENSEA, CNRS, 95000 Cergy-Pontoise, France
5) ASAsense cvba, Bruges, Belgium

Summary
Relying on monitoring networks to compute or improve noise maps is an increasingly used approach. To be able to use this approach to provide adequate temporal treatments, a good understanding of the temporal variations within urban sound level time series is required. This paper provides an in-depth statistical analysis of the temporal characteristics of urban sound environments, on the basis of a wide measurement campaign during 8 month, at 23 measurement stations in Paris, which cover a large variety of urban sound environments. The time series of sound levels were recorded continuously with a 125ms-time resolution, from which $L_{A50,1h}$ values were extracted. In total, 72 time-slots of interest are defined (24 1h-periods covering all days of the week). The statistical analysis determines for each station the Daily Average Noise Pattern (DANP), and for each of the 72 time-slots the 1h-Generalized Extreme Values distributions. The Generalized Extreme Values distributions are found to outperform the normal distributions to model the $L_{A50,1h}$ distributions. In addition, the average sound level differences between these 72 1h-time periods are calculated along with their variability, resulting in 72×72 delta matrices that describe the temporal relations between sound levels. This database is then used to develop two models, which aim to estimate DANP based on a limited amount of measurements. The model M1 relies on the delta matrices, whereas the model M2 consists of a weighted average of the DANP that are stored in the database in which the weights are based upon measures of similarity between the stations. Both models rely on probability density functions, and provide a measure for the reliability of the estimated noise levels. A test of both modelling approaches through simulated measurements shows that the model M1 seems to be more robust in case measurements are inaccurate. Beyond these two models, the proposed database could serve in the development of further models that aim to estimate sound levels based on a limited amount of measurements.

PACS no. 43.50.Rq, 43.50.Yw

1. Introduction

The Directive 2002/49/EC stands as a response to city dwellers preoccupations regarding noise. It requires that European cities of more than 100 000 inhabitants elaborate and broadcast strategic noise maps [1]. These maps present Lden values, which correspond to the energetic average sound level of the environment with a 5 dB and 10 dB penalty for evening and night periods respectively. However, the temporal dynamics of the sound levels also plays a role in the emergence of annoyance, which is influenced both by the fast dynamics (peaks of noise, rhythm imposed by the traffic lights, etc.) and the slow dynamics (city morning awakenings, issues with too high leisure noise levels, etc.) in the sound level [2, 3, 4].

Simulated maps were historically the preferred method to produce the Lden strategic noise maps. They combine source emission and sound propagation calculations, and may be performed through a large variety of software dedicated to sound mapping [7, 8]. Simulated maps have been progressively completed with noise observatories, which record the sound level time series at strategic lo-
cations, through high-quality sound level meters [9]. Both approaches suffer however from some limitations: the former only considers a limited amount of sources, whereas the latter is expensive and limited in spatial coverage due to the difficulty to interpolate measured levels [10].

New technological solutions arose, which enable collecting plethora of noise data in urban area. They rely either on dense low-cost sensor network deployments [11], or on participative data collection via smartphone applications [12, 13, 14, 15, 16, 17]. Such solutions permit a much larger spatial coverage as compared to the use of high-quality sound level meters, but face in counterpart some methodological and metrological issues [18, 19].

Beyond relying on simulated maps or measurements, a third approach that merges the two first ones within a common modelling framework has been developed, with the perspective to converge towards more accurate maps. The objective here is to correct local sound levels based on measurements through data fusion techniques [20], or to continuously modify through measurements the modelling parameters [21]. However, some questions remain unresolved, concerning the indicators to produce [22], or how to efficiently account for the spatial and temporal vari-ability of noise. Sound environments are indeed characterized by their very pronounced short-term variations [23], and their daily and weekly periodicity [24]. Another difficulty for building such a modelling framework stands in the lack of reference data for validating the constructed models or associating the produced maps with a level of confidence or an uncertainty.

In this paper, the main specificities of urban sound level time series are extracted from a detailed analysis of 8 months of measurements collected in 23 points in Paris, France. The constituted database gathers the information required to build temporal sound level interpolations. A possible use of the database is illustrated through the proposal of a modelling framework that estimates temporal trends in daily average sound levels, and sound level probability density functions, based on a limited amount of measurements, by using the statistical properties of the sound level times series within the database.

2. Data collection

2.1. Measurement stations

Noise data collection was performed at 23 long-term monitoring stations, during 8 months lasting approximately from July 2014 to February 2015 (see the measurement periods in Table 1). The measurement devices consisted of an ALIX 3D3 single-board computer, an industrial grade 8 GByte Compact-Flash card, a Knowles microphone with 3D-printed holder and rain screen, and a windscreen with a diameter of 9 cm. Each of the measurement devices was calibrated using a B&K 4231 calibrator, and subsequently their accuracy was measured in an anechoic room. The 125ms-sound pressure levels $L_{Aeq,125ms}$ were collected continuously, from which sound indicators were calculated with a 1h time-resolution. The choice of this time-resolution is supported by the facts that this period is often encountered in the literature to characterize sound environments [25], and that the stability of sound environments at this time-scale has been shown in [26]. Shorter time-resolutions are discussed in Section 6. The 23 stations are distributed in the 13rd district of Paris, within an area of about 4 km$^2$, as depicted in Figure 1. The stations cover different road traffic and morphologic configurations, with low to high traffic volumes, pedestrian streets, and some points are located near parks. Stations were distributed within the 1st and 5th floor. The impact of the positioning of the long-term stations on the proposed modelling approach is discussed in Section 6.

2.2. Sound indicators

Among the 1h resolution calculated indicators, the study only focuses on $L_{A50}$, which is of particular interest for estimating sound pleasantness in urban area [27, 28]. However, the constituted database also includes other indicators, such as $L_{Aeq}$, the statistical indicators $L_{A90}$, $L_{A50}$, $L_{A10}$ and $L_{A1}$, and advanced indicators highlighting the short-term temporal dynamics (Mask Index, number of events [22]).

A second derived indicator is the DANP$_{i,s}$ at each station $s$, where $i$ stands for the day-of-the-week, $i = \{mf; sat; sun\}$. Three typical days-of-the-week are considered: “Monday-to-Friday” (mf), “Saturday” (sat) and “Sunday” (sun), which are known to show different temporal trends in the sound level [24][25]. The DANP$_{i,s}$ consists of a series of 24 $L_{A50,h,i,s}$ values, where $h$ stands as the beginning of the 1h time-period, $h = \{0; 1; \ldots; 23\}$, representative of the temporal period of interest. In this study this period is 8 months, but practically it could be one or several years (thus seasonal sound level variations are not fully included in this modelling). These $L_{A50,h,i,s}$ values are the averaged value of the $L_{A50,h,i}$ values calculated during $i$ and $h$ over the sampling period (for instance the $L_{A50,8h, sat, P2}$ is the average of the $L_{A50,1h}$ values calculated at the station P2 on Saturdays in the 1h-period [8-9]h). In the following of the paper, to alleviate notations $L_{A50}$ will correspond to $L_{A50,1h}$ values unless otherwise specified. Thus a given $L_{A50,h,i,s}$ value calculated during a random period...
sample at hour $h$ and day-type $i$ can significantly differ from the $L_{A50,h,i,s}$ value, because of the high sound level temporal variability, as will be shown in Section 3.3. So, in this paper the main interest is not to estimate dynamically the $L_{A50,h,i,s}$ time series at a given location $s$, but to estimate the $L_{A50,h,i,s}$ values, because: (i) $L_{A50,h,i,s}$ and the $L_{A50,h,i,s}$ values mainly differ by sound level variability that is due to processes that are random by nature (traffic variability, presence of very noisy vehicles, etc.), which makes the individual values less representative of the sound environment at a given location, (ii) $L_{A50,h,i,s}$ values correspond more to the noise mapping goals recommended in the Environmental Noise Directive (END).

3. Data analysis

3.1. Analysis of the daily average noise patterns

3.1.1. Description of the daily average noise patterns

A statistical analysis of the 8 months of collected data is presented in this section, with the aim to underline the specificities of the urban sound environments that must be taken into account for temporal sound level interpolation. The DANPs, deduced at each location $s$ from the measurements at the 23 stations are displayed in Figure 2. The Figure 2 highlights the large amplitude of sound levels in the area, which have a 20 dB range between the noisiest point P3 and the quietest one P9. This variety of sound environments can also be illustrated by the fact that the $L_{A50,h,i,s}$ values at the loudest points at night are higher than that the $L_{A50,h,i,s}$ values at the quietest points during the day. This shows the large variety of the encountered urban sound environments even at a small spatial scale. Despite this large range in sound levels, the DANP are highly correlated, mainly because of the night and day sound levels alternation.

3.1.2. Correlations between the daily average noise patterns

The correlations between the sound level time series are presented in Figure 3 (each element of the matrix represents the Pearson correlation between the 24 $L_{A50,h,i,s}$ val-
ues for a given couple of stations $s_1$ and $s_2$). The sound level temporal trends are in general highly correlated, with average Pearson correlation coefficient $\bar{r}$ (average of the $23 \times 23$ coefficients per matrix) of 0.95, 0.93 and 0.90 for $mf$, $sat$, and $sun$, respectively. This advocates estimating the temporal trends in the sound level at a given location based on observations gathered at long-term stations. Figure 3 reveals in addition some specificity in the domain of encountered sound level time series at certain points. In particular, the sound level evolution at P2 is poorly correlated to the other ones, showing high $L_{A50,h,s=0}$ values in the $sat$ and $sun$ nights ($L_{sat} = 0.58$ and $L_{sun} = 0.69$). These high levels are explained by the particular location of P2, which is situated in “la Butte aux Cailles” (see Figure 1), a street with leisure activities such as bars and restaurants that generate high noise levels at these evening periods. P3 also shows low $L_{A50,h,s=0}$ correlations with the other points of the dataset ($L_{sun} = 0.75$). This is due to the slower Sunday morning noise levels increase, between 10h and 13h, relatively to the rest of the points (despite the fact that sound levels are higher).

Thus the correlation matrix shows that relying on a database that hosts general statistics on temporal trends in the sound level proves useful in most of the cases for deducing information at new locations, but may be more difficult at locations whose sound environments are potentially untypical. This difficulty will be evaluated and discussed in Section 5.

3.1.3. Linear relations between the daily average sound level patterns

The high correlations between the daily average sound level patterns suggest the possibility to link them through the following linear regressions $DANP_{i,s} = a_{i,s_1,s_2} \cdot DANP_{i,s_1,s_2} + b_i$, with $s_1$ and $s_2$ two monitoring stations. The interest of such modelling is to stretch the DANP through the $a_{i,s_1,s_2}$ parameter, and shift it through the $b_i$ parameter, and thus estimate $DANP_{i,s}$ based on $DANP_{s_1,s_2}$. Because of these high correlations, and because the sound level amplitude is found to be limited to a range of about 20 dB in the area, the domain of the $\{a, b\}$ values that link two given stations is restrained. Figure 4 shows the domain of encountered $\{a, b\}$ values for each combination of two of the 23 valid long-term stations, that is $23 	imes 22 = 506$ combinations. Noticeably, the crosses with $a > 2.5$ all correspond to the point $P_{21}$ (in magenta in Figure 4), which shows a very low night-to-day sound levels amplitude, of 9.6 dB whereas it is on average 14.6 dB for the other stations (see Figure 2). Equations are proposed to delimit the expected linear regressions between a couple $DANP_{i,s_1}$ and $DANP_{i,s_2}$ from the database. The domain of $\{a, b\}$ values, represented in Figure 4 by dotted red lines, can be restrained to the following domain:

$$ \begin{cases} b & = 80 - 47.5a, \\ b & = 80 - 62.5a, \\ 0.1 < a < 4. \end{cases} $$

In addition, Figure 4d presents the Root Mean Square Errors (RMSE) of the $L_{A50}$ estimates that are associated to these linear regressions. The errors are remarkably low, with average RMSE values of 2.6 dB, 1.9 dB and 2.2 dB for $mf$, $sat$ and $sun$, respectively, and 73%, 91% and 80% of RMSE values below 3 dB for $mf$, $sat$ and $sun$, respectively. This proves the interest to link DANP though linear regressions. The fact that the RMSE values are slightly higher for $mf$ despite higher correlations between the DANP can be explained by the higher sound level amplitude for $mf$, compared with $sat$ and $sun$. Finally, Figure 4d highlights the small proportion of high errors relative to the sat regressions, with 5.5% of the RMSE values exceeding 10 dB; these errors correspond to the point
P2, whose low correlations with the other stations due to leisure activities have been discussed in Section 3.1.2. 

3.2. Relations between sound levels during different time periods

The repeatability of the daily average sound level patterns suggests the possibility to estimate the DANP \(_{f,s}\) based on samplings of a few \(L_{A50,s}\) values at 11:00, assuming that a measurement achieved for instance on a Tuesday at 11:00 informs about the sound levels that can be expected on a Saturday at 16:00. Therefore, nine matrices \(\delta_{i_1,i_2}\) are determined, with \(i_1\) and \(i_2 = \{mf, sat, sun\}\), each of the 9 matrices being of size \(24 \times 24\), and each of its elements \(\delta_{i_1,i_2}(h_1,h_2)\) containing the estimated delta value \(\delta_{i_1,i_2}(h_1,h_2) = L_{A50,h_1,i_1} - L_{A50,h_2,i_2}\).

In addition, these delta matrices are associated with a given standard deviation, expecting that a \(L_{A50,h,i}\) value collected for instance on a Tuesday at 11:00 tells more about the sound levels on a Tuesday at 10:00 than on a Sunday at 05:00. To account for this standard deviation, each of the nine matrices \(\delta_{i_1,i_2}\) is associated with a matrix \(u_{i_1,i_2}\) that gathers the standard deviation associated to the delta value. For example, \(\delta_{mf,mf}(10h,11h) = -0.3\) dB and \(u_{mf,mf}(10h,11h) = 3.2\) dB, whereas \(\delta_{mf,sun}(10h,05h) = -11.2\) dB and \(u_{mf,sun}(10h,05h) = 4.1\) dB. In practice, these \(\delta_{i_1,i_2}\) and \(u_{i_1,i_2}\) matrices are determined from the \(L_{A50,h,i,s}\) values collected over the 34 weeks at the 23 valid stations. At each station, for a given couple \{\(i_1, h_1\)\} and \{\(i_2, h_2\)\} during the same month (in order to avoid seasonal noise level variations), one obtains \(4^4 = 16\) \(L_{A50,h_1,i_1} - L_{A50,h_2,i_2}\) values per month if \(i_1\) and \(i_2 = sat\) or \(sun\) (supposing there is exactly four weeks in the month), \(16 \times 5 = 80\) values if \(i_1\) or \(i_2\) equals sat or sun and the other equals \(mf\) (5 days from Monday to Friday), and \(16 \times 25 = 400\) values if \(i_1\) and \(i_2 = mf\). Over the 8 months of measurements and the 23 stations, this corresponds to a maximum number of values of 2944 if \(i_1\) and \(i_2 = sat\) or \(sun\), 14720 if \(i_1 = mf\) and \(i_2 = sat\) or \(sun\) or the opposite, and 73600 if \(i_1\) and \(i_2 = mf\). Figure 5 illustrates some typical \(\delta_{i_1,i_2}(h_1,h_2)\) profiles. The profiles follow unsurprisingly the DANP shapes, with \(\delta\) values ranging between \(-15\) and \(+15\) dB according to the \{\(h_1, h_2\)\} couples. The diagonal of the \(\delta_{i_1,i_2}\) matrix is null, but the diagonal of the \(u_{i_1,i_2}\) matrix is not; for instance, \(u_{mf,mf}(10h,10h) = 3.2\) dB. This is due to the inter-day \(L_{A50}\) variability. In addition, Figure 5 shows the increased standard deviation for couples of time-of-the-day values that are distant, or when the couple is composed of different days-of-the-week. These standard deviations are depicted in Figure 6, which represents a matrix of all the \(u\) values, which evolve between 1.9 and 5.2 dB. The Figure 6 highlights the lack of representativeness of measurements achieved during night time periods for estimating day sound levels, and vice versa. It also underlines the stability of sound level variations on Sunday, which are the time periods associated with the lowest standard deviations. This advocates for measurement strategies that include samplings during both day and night periods, and more generally a large variety of periods.

3.3. Sound level distributions during different periods

This section aims to describe the sound level variability at each period. This will help to understand the shape of the temporal distributions of sound levels, and to associate a standard deviation to measurements. The distributions of the \(L_{A50}\) values are calculated at each of the 23 valid sta-
distributions, and are described by

\[ y = f(x|\kappa, \mu, \sigma) \]

\[ = \frac{1}{\sigma} \exp \left[ - \left( 1 + \kappa \frac{x - \mu}{\sigma} \right)^{-1/\kappa} \left( 1 + \kappa \frac{x - \mu}{\sigma} \right)^{-1-1/\kappa} \right]. \] (1)

where \( \mu \) is the location parameter, \( \sigma \) is the scale parameter, and \( \kappa \) is the shape parameter. GEV have been successfully used in the past in the field of environmental researches, to estimate air pollution [29, 30] or precipitation [31] extreme episodes. It is expected that its ability to reproduce asymmetry or long tails in the distributions responds to noise requirements, whose \( L_{A50} \) values are bounded approximately between 40 and 70 dB(A) (see Figure 2) and in which extreme episodes are also expected to occur.

Normal distributions and GEV distributions have been fitted on the \( L_{A50,1h} \) distributions (23*72 distributions of \( L_{A50,1h} \) values in total). The relevance of both distributions is tested through a Two-sample Kolmogorov-Smirnov test (ktest2 function in Matlab), which tests the null hypothesis that data in vectors \( x1 \) and \( x2 \) comes from populations with the same distribution. The GEV distribution outperformed the normal distribution to model the 23*72 \( L_{A50,1h} \) distributions over the 8 months of collected data. Indeed, the Two-sample Kolmogorov-Smirnov test rejected the null-hypothesis that the 23*72 \( L_{A50} \) distributions follow a normal distribution in 15.6% of the cases at a 5% significance level (averaged p-value > 0.05), and for only 8.2% of the cases for the GEV distribution (5% significance level, averaged p-value > 0.05). Indeed, GEV distributions reproduce better both the sound level distributions that are often asymmetric, and the very high sound levels that can be assimilated to rare events.

As shown in Figure 7 for the station P1, the GEV distribution is sensible to the shape of the distribution through the \( k \) parameter, showing generally a longer tail on the right for low noise levels, and a longer tail on the left for high noise levels. Sound level distributions depend on the day of the week, showing a clear distinction between Monday-to-Friday, Saturday and Sunday periods, as already shown in [25]: sound levels are higher from Monday to Friday during the day periods, but become higher during the night periods. The GEV distribution parameters \( \mu_{h,i,s}, \sigma_{h,i,s} \) and \( \kappa_{h,i,s} \) are stored for each time-of-the-day \( h \), each day-of-the-week \( i \) and each long-term station \( s \), resulting in 72 values per station and
are not chosen completely randomly, but vary instead according to the time of the day: the probability to perform a measurement during night time periods is fixed to be 10 times lower than during day time periods. In addition, each of the $K$ measures is associated to a probability density function (pdf), $L_{A50,h(i),j}(k)$, with a standard deviation of $\sigma_{k,h(j),S} = \sqrt{(\sigma_k + \sigma_{h(i),S})^2}$.

4. Presentation of the Model $M_1$

The model $M_1$ relies both on the $K$ $L_{A50,h(i),j}(k)$ measurements at $S$ and the delta matrices $\delta_{i,j}(h_1, h_2)$, in order to estimate the DANP$_{M_1}$, First, each of the $K$ $L_{A50,h(i),j}(k)$ estimated values is extended to 72 probability density functions $p_{M_1,k,h,i}(L_{A50})$, with $h = \{1; \ldots; 24\}$ and $i = \{m,f,sat,sun\}$, centred at $L_{A50,h(i),j}(k) + \delta_{i,j}(h_1, h_2)$, and with the standard deviation $\sigma_{k,h(i)}$.

Then, at each 1h-time frame $\{h,i\}$, the K $p_{M_1,k,h,i}(L_{A50})$, are averaged to constitute the $p_{M_1,k,h,i}(L_{A50})$. The estimated DANP$_{M_1}$ is finally the centre of gravity of the pdf $p_{M_1,k,h,i}(L_{A50})$. Thus, the model $M_1$ gives by construction more weight to accurate measurements, which result in distributions with a smaller dispersion. In addition, the model $M_1$ results theoretically in an estimated DANP$_{M_1}$ that is more accurate at the time frames with many measurements, since the $\sigma_{i,j}(h_1, h_2)$ value is minimum when $h_1 = h_2$.

4.2. Presentation of the Model $M_2$

The model $M_2$ relies both on the $K$ $L_{A50,h(i),j}(k)$ measurements at $S$ and the DANP$_{S}$ of the stations $S$ stored in the database to estimate the DANP$_{M_2}$. First, The $K$ $L_{A50,h(i),j}(k)$ values are used to estimate the similarity $w_{S,S}$ within $S$ and each monitoring station $S$ from the database. The similarity $w_{S,S}$ is expressed as the inverse of the averaged Euclidean distance between the $K$ $L_{A50,h(i),j}(k)$ values and the corresponding $L_{A50,h(i),j}(k)$ values at $S$,

$$w_{S,S} = \frac{1}{\sqrt{\sum_K \left( L_{A50,h(i),j}(k) - L_{A50,h(i),j}(k) \right)^2}}.$$

These similarities then serve to estimate DANP$_{M_2}$ as a linear combination of the DANP$_{S}$ values, weighted by the $w_{S,S}$ values, such as

$$\text{DANP}_{p,m} = \sum_S w_{S,S} \text{DANP}_{S}.$$

Probability density functions $p_{w,h,i}(L_{A50})$ for each $h$ and $i$ are similarly expressed as the average of the GEV distributions at each station $S$, weighted by the $w_{S,S}$ values. Thus, the formed DANP$_{w,i}$ is mostly influenced by the stations whose sound level evolution is the most similar to $S$.

Then, the coefficients $a$ and $b$ of the linear regression that links the DANP$_{w,i}$ to the $K$ measured values $L_{A50,h(i),j}(k)$ are determined. This linear regression aims to correct for the difference between the DANP at $S$ and
Figure 8. (Colour online) DANPM₁ and DANPM₂ estimates at P1, with \( \sigma_k = 1 \) dB. Left: \( K = 5 \) measurements. Right: \( K = 20 \) measurements. The colorbar represents the probability density function.

the DANPs that are stored in the database, both in terms of sound level mean values and amplitude (day levels minus night levels). In practice, the parameters \( a \) and \( b \) are calculated based on the measurements collected during \( H(K) \) and the DANP \( \mu_k(H(K)) \) values at the same time periods. The domain of research for \( a, b \) is limited by the functions given in Section 3.1.3.

The parameters \( a \) and \( b \) are used to shift the centre of gravity of the \( p_{h,k}(L_{A50}) \) probability density functions, to constitute the new pdf, \( p_{r,h}(L_{A50}) \) which hopefully better reflects the sound level at \( S \). Finally, the \( p_{M_{h,k}}(L_{A50}) \) is at each \( \{h, i\} \) the average between the \( p_{r,h}(L_{A50}) \) and the measurements \( p_{k,h(k)=h, i(k)=i}(L_{A50}) \), which are achieved at \( \{h, i\} \).

Thus, the model \( M₂ \) gives by construction an important weight to measurements at the time frames where numerous measurements are collected, since the measures only impact in \( M₂ \) estimates at their precise time frame, contrarily to the model \( M₁ \). In addition, when the number of measurements is very limited, the use of the linear regression guarantees in theory coherent estimated sound level evolutions.

4.4. Illustration of the modelling frameworks

4.4.1. Case 1: \( \sigma_k = 1 \) dB

In Case 1, measurements are achieved at P1 with a class-1 sound level meter and a professional operator that follows rigorously the measurement protocol, resulting in a very low measurement standard deviation \( \sigma_k = 1 \) dB. The pdf functions and the DANP estimated with \( M₁ \) and \( M₂ \) are illustrated in Figure 8, for \( K = 5 \) measurements (left) and \( K = 20 \) measurements (right), along with the real DANP value at P1 DANP_{real} and the DANP_{conc}, which is the average of the 22 DANP.

First, this very low standard deviation of \( \sigma_k = 1 \) dB does not guarantee that measurements estimate the \( L_{A50} \) values with the same accuracy of 1 dB. In this example, the 5 \( L_{A50} \) values derived from measurements stand within a range of about 3 dB around the actual \( L_{A50} \) values, which result from the \( L_{A50} \) variability at a given time-of-the-day that was described in Section 3.3. This advocates for sampling strategies that cover a high number of time frames. Nevertheless, both models \( M₁ \) and \( M₂ \) succeed, even with \( K = 5 \), in estimating precisely the DANP. The shape of the noise level temporal evolution is reproduced with a good reliability even in the case when no measurements during the night periods are available, thanks to the high temporal correlation of the urban sound levels that was explained in Section 3.1.2, on which both models are based. In addition, the sound levels are estimated with a good accuracy: they are higher in P1 than in the average of the other monitoring stations by about 5 dB. Sound levels are however slightly underestimated, because simulated measurements were achieved during periods when sound levels were below the usual values (compare in Figure 8 the black crosses and the black curve). The increase in the number of measurements solves this default as it increases the representativeness of the sampling; as a result, sound levels are estimated with a good accuracy by both models when \( K = 20 \).

Moreover, the case with \( K = 20 \) in this example underlines the specificities of both models. In the mf period, the model \( M₂ \) seems to improve the noise level estimation during the night period, perhaps because the measurements help estimating the sound levels in \( M₂ \) through linear regression. Inversely, the measurements taken during Sundays lead to sound levels that are too much amplified by the linear regression, resulting in underestimated sound levels at night. In counterpart, the model \( M₂ \) better describes the high sound levels at P1 during Sundays around
Figure 9. (Colour online) DANP\(_{M_1}\) and DANP\(_{M_2}\) estimations at P1, with \(\sigma_k = 10\, \text{dB}\). Left: \(K = 5\) measurements. Right: \(K = 20\) measurements.

12:00, which are a particularity of the sound level evolution at P1. This can be explained by the fact that sound level values at their time-period are specifically taken into account by the model.

4.4.2. Case 2: \(\sigma_k = 10\, \text{dB}\)

In Case 2, measurements are obtained at P1 with a standard deviation of \(\sigma_k = 10\, \text{dB}\). This corresponds to very inaccurate measurements, such as provided by participative measurements, because of very short sampling durations, protocol issues or apparatus malfunctions. As a consequence, the estimated probability density functions are much wider, revealing the unreliability of estimates. None of models \(M_1\) and \(M_2\) converge towards accurate DANP when \(K = 5\). In this example, the 5 measured \(L_{A50}\) values highly underestimate the actual ones, resulting in an underestimated DANP with both \(M_1\) and \(M_2\). However, the shape of the DANP is accurately reproduced by the models despite the low number of measurements and their inaccuracy, showing the interest of the modelling approaches.

The DANP at P1 is better approached with \(K = 20\), as the increased number of measurements makes the average of the measured sound levels converge towards their actual value despite the individual measurements errors, assuming that there are no systematic biases in the measurements, or that they are known and taken in to account. The case with \(\sigma_k = 10\, \text{dB}\) shows the limitation of the model \(M_2\) when individual measurements are imprecise. Indeed, some periods are associated to a very high error, because the model \(M_2\) gives an important weight to each individual measurement. It can be helpful to highlight the sound level specificities at a point for a given time period (for example when a period shows abnormally high sound levels because of human activities), but this approach is risky when the individual measurements on which it relies are imprecise. Section 5 will compare the benefits of both models over the total set of points through a leave-one-out approach.

5. Results

The models \(M_1\) and \(M_2\) are run successively at the 23 points, excluding each time the station of interest from the database. Models \(M_1\) and \(M_2\) are tested on their ability to reconstruct the weekly average noise pattern \(WANP\) at each point \(s\), which is made of the concatenation of 5 DANP\(_{mf}\), one DANP\(_{sat}\) and one DANP\(_{sun}\), thus representing the one-week average noise pattern. Two indicators of quality are defined to evaluate the models:

- The RMSE\(_{WANP}\) measures the difference between the estimated and the actual \(WANP\). It consists of the RMSE between the 168 \(L_{A50,h,i,s}\) values that form the estimated and the actual \(WANP\);
- the \(\delta_{WANP}\), which is difference between the estimated and the actual \(WANP\), which are the arithmetical averages of the 168 \(L_{A50,h,i,s}\) values that form the estimated and the actual \(WANP\).

These indicators are averaged over these 23 runs and depicted in Figure 10, along with their standard deviation, for the three different measurement standard deviation values \(\sigma_k = 1, 5\) and \(10\, \text{dB}\), according to the number of measurements.

In the case of very accurate individual measurements, when \(\sigma_k = 1\, \text{dB}\), a very low number of measures guarantees accurate \(WANP\) estimates, with RMSE\(_{WANP}\) and \(\delta_{WANP}\) values below 3 dB even with one measure. This is due to the strong temporal repeatability of the sound level patterns, 3 dB corresponding approximately to the dispersion of the \(L_{A50}\) values during the experiment.

If the results of the two models are almost similar for the \(WANP\) estimates, the \(WANP\) are reproduced with a better accuracy by the model \(M_1\), especially when the standard deviation associated to the individual measurements...
The statistical analysis resulted in the determination, for each of the 23 locations in Paris during 8 months, of the temporal structure of sound levels, the model M1 permits a very precise estimation of both the noise levels and their temporal evolution, even with inaccurate measurements. As an example, with the model M1 and with \( \sigma_k = 10 \text{ dB} \), \( K = 10 \) measurements are sufficient to estimate \( \text{WANP} \) with an error of about 3 dB (\( \delta_{\text{WANP}} \) from 1 to 5 dB according to the points), and 15 measurements are sufficient to estimate the WANP with a RMSE\(_{\text{WANP}}\) of 3 dB.

### 6. Discussion

This paper provided an in-depth statistical analysis of the temporal characteristics of the sound level in urban environments, on the basis of a wide measurement campaign at 23 locations in Paris during 8 months. The time series of sound levels were recorded continuously with a 125 ms-time resolution, from which \( L_{\text{A50,1h}} \) values were extracted. The statistical analysis resulted in the determination, for each of the 23 point and each “Monday-to-Friday”, “Saturday” and “Sunday” 1h-time periods (that is 72 time periods), of the \( L_{\text{A50}} \) distributions, through the form of Generalized Extreme Values distributions. In addition, the average \( L_{\text{A50}} \) differences between these 72 1h-time periods are calculated along with their variability, resulting in 72*72 delta matrices that describe the temporal relations between \( L_{\text{A50,1h}} \) values. The results from this wide statistical analysis was then used to build two models that estimate Daily Average Noise Patterns at a given location, relying on measurements at other locations and the information stored in a database.

The model M1 relies on the measurements along with the delta matrices, which help estimating DANP although not all time-periods are covered by measurements. The model M2 relies instead on the DANP stored in the database, which are deformed thanks to the measurements through linear regressions, adding \( \text{a posteriori} \) the realized measurements to finally estimate the DANP. Both models rely on probability density functions, thus accounting for the reliability of the estimated sound levels. A test of both modelling approaches through simulated measurements shows that the model M1 seems to be more robust in case measurements are inaccurate, which is typical for participative measurement schemes.

The study has some limitations:
- This study is the limited set of 23 observed sound environments. The domain of validity of the proposed models is restrained to the variety of the observed sound environments. However, the similarities between the sound levels temporal trends are probably high from one city to the other, making it \( \text{a priori} \) possible to use the dataset for other cities. Comparisons between measurement campaigns in various cities are nevertheless required to test this hypothesis. In addition, the constituted database is meant to be enriched in the future with
any new long-term measurement associated with the proposed statistical analysis, including measurements collected in various cities. One expects that monitoring stations at locations with similar morphologies or traffic situations but from different cities, will prove useful to apply the proposed methodology at new locations.

- The difference in configurations between the monitoring stations is not handled in the study. Differences in measurement heights are for instance likely to affect sound levels dynamics. Researches are required to decide how different measurement configurations can be integrated within a common noise monitoring network, for instance through transfer functions. However, these differences do not compromise the proposed approach, as the correlations between the daily average noise patterns are high; the robustness of the approach to these different measurement configurations has however to be verified for other indicators, such as $L_{A10}$.

- The approach relies on 1h-time periods. The estimation of $L_{A50}$ values with shorter time-resolutions could imply higher errors. However, the stability of noise levels at the 1h-scale limits the range of errors: statistics calculated from the DANP presented in Figure 2 show that the average difference between two consecutive $L_{A50,1h}$ values of the DANP is 1.3 dB(A), and its maximum value 5 dB(A). This however suggests the possible higher errors with short time-periods in the morning (periods from 6h to 8h), when the sound level variations are high.

Other models could be proposed based on the same database. The measure of similarity included in model $M_2$, which simply relies on a calculation of RMSE, could be based on similarity evaluations that call for both relevance and redundancy metrics. Furthermore, specific outlier detection algorithms could be designed, to exclude the abnormal measured $L_{A50}$ values, such as proposed in [19]. The difficulty then stands in the need to exclude default measures but still capturing the specificities in noise level evolution (periods with atypical sound levels).

Beyond the two proposed models, the database of noise level characteristics used in this paper could serve to build new modelling frameworks dedicated to the evaluation of time series of sound levels based on sparse measurements. This research will then contribute to an increased understanding and characterization of urban sound environments through monitoring networks.

Acknowledgements

This work was performed in the framework of the GRAFIC project (grant number 1317C0028), funded by ADEME (French Environment and Energy Management Agency).

References


[19] A. Can, G. Guillaume, J. Picaut: Cross-calibration of par-
ticipatory sensor networks for environmental noise map-
[21] W. Wei, T. Van Renterghem, B. De Coensel, D. Bottel-
dooren: Dynamic noise mapping: A map-based interpola-
tion between noise measurements with high temporal reso-
[22] A. Can, G. Guillaume, B. Gauvreau: Noise indicators to di-
[23] A. J. Torija, D. P. Ruiz, A. Ramos-Ridao: Required sta-
proaches to predict urban street noise levels using fixed and temporary microphones. Journal of Environmental Moni-
[26] L. Brocolini, C. Lavandier, M. Quo, C. Ribeiro: Mea-
[27] P. Aumond, A. Can, C. Lavandier, B. De Coensel, D. Bot-
[28] P. Ricciardi, P. Delaitre, C. Lavandier, F. Torchia, P. Au-
[31] Z. Rulfová, A. Buishand, M. Roth, J. Kyselý: A two-
component generalized extreme value distribution for precip-