Probabilistic modelling of the temporal variability of urban sound levels

Authors : Arnaud Can[1], Pierre Aumond[1], Bert De Coensel[2], Carlos Ribeiro[3], Dick Botteldooren[2], Catherine Lavandier[4]

[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), F-44341 Bouguenais, France

[2] Acoustics Group, Department of Information Technology, Ghent University, Technologiepark-Zwijnaarde 15, 9052 GENT, Belgium

[3]Bruitparif, France, 90-92 Avenue du Général Leclerc, 93500 Pantin
[4]ETIS - UMR 8051, Université de Cergy-Pontoise St-Martin, 2 avenue Adolphe Chauvin, 95032 Cergy-Pontoise Cedex

Keywords : urban sound environment, sampling strategies, daily average noise pattern, probabilistic modelling,

Highlights:

- **A wide acoustical measurement campaign is realized**
- **A database of urban sound environment characteristics is proposed**
- **Daily average noise patterns are extrapolated from short term measurements**

Abstract

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 M_1 relies on the delta matrices, whereas the model M_2 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 M_1 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.

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 L_{den} 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 to high leisure noise levels, etc.) in the sound level [2][3][4].

Simulated maps were historically the preferred method to produce the L_{den} 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 locations, 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 variability 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 where 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 km², 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.

Table 1. Description of measurement locations and time periods

Point	Address	Floor	Date begin	Date end	Description of the sound environment
P_1	69 Boulevard Auguste Blanqui	1st	04/07/2014	28/02/2015	Boulevard with dense traffic and aerial metro

P_2	3 rue de la Butte aux Cailles	2nd	08/07/2014	28/02/2015	Small street with leisure activities (bars, restaurants)
P_3	138 avenue d'Italie	2nd	01/07/2014	28/02/2015	Street with dense traffic
P_4	45 Rue Croulebarbe	5th	02/07/2014	28/02/2015	Small street, vicinity of a small park
P_5	25 bd Arago	2nd	08/07/2014	28/02/2015	Boulevard with dense traffic
P_6	141 rue du Château des Rentiers	2nd	09/07/2014	28/02/2015	Residential streets
P_7	58 avenue de Choisy	5th	15/07/2014	28/02/2015	Boulevard with dense traffic
P_8	5 rue Wurtz	4th	17/07/2014	28/02/2015	Very small street
P_9	10 passage Barrault	1st	02/07/2014	28/02/2015	Very small street
P_{10}	185 bd Vincent Auriol - Esc. 34	3rd	17/07/2014	28/02/2015	Boulevard with dense traffic
P_{11}	5 rue Philibert Lucot	5th	14/07/2014	28/02/2015	Small street, vicinity of a dense street
P_{12}	180 avenue de Choisy	1st	16/07/2014	28/02/2015	Boulevard with dense traffic, vicinity of a very dense roundabout
P_{13}	76 rue Barrault	3rd	04/07/2014	20/02/2015	Very small street, residential neighbourhood
P_{14}	78-84 rue Brillat Savarin	5th	26/06/2014	28/02/2015	Very small street, residential neighbourhood
P_{15}	51 avenue des Gobelins	2nd	03/07/2014	28/02/2015	Street with dense traffic
P_{16}	29 Place Jeanne d'Arc	4th	02/07/2014	12/02/2015	Street with dense traffic
P_{17}	26 rue de Rungis	4th	04/07/2014	20/02/2015	Small street
P_{18}	11bis rue de l'Amiral Mouchez	2nd	07/07/2014	15/02/2015	Small street, vicinity of a large intersection
P_{19}	37 rue Albert	2nd	03/07/2014	28/02/2015	Very small street, residential neighbourhood
P_{20}	137 avenue de Choisy	3rd	07/07/2014	28/02/2015	Boulevard with dense traffic
P_{21}	19 rue Godefroy	1st	02/07/2014	28/02/2015	Vicinity of a very dense roundabout
P_{22}	198 rue de Tolbiac	4th	27/06/2014	28/02/2015	Street with dense traffic
P_{23}	140 rue Léon Maurice Nordmann	1st	22/05/2014	28/02/2015	Small street

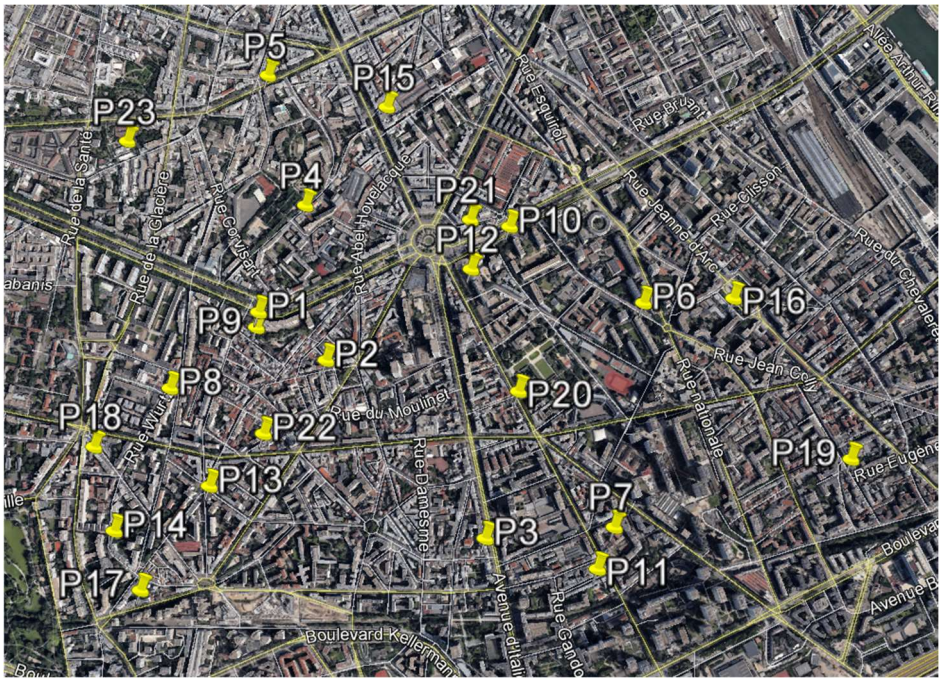

Figure 1. Location of the 23 long-term measurement stations.

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 $\overline{L_{A50,h,i,s}}$ values, where *h* stands as the beginning of the 1h time-period, *h* = {0; 1; …; 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 $\overline{L_{A50,h,i,s}}$ values are the averaged value of the $L_{50,h,i,s}$ values calculated during *i* and *h* over the sampling period (for instance the $\overline{L_{A50,8h,sat,P2}}$ is the average of the $L_{A50,1h}$ values calculated at the station P_2 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 sample at hour *h* and day-type *i* can significantly differ

from the $\overline{\mathrm{L}_{\mathrm{A50},h,i,s}}$ 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 $L_{A50,s}$ time series at a given location *s*, but to estimate the $\overline{\mathrm{L}_{\mathrm{A50},h,i,s}}$ values, because: (i) $\overline{\mathrm{L}_{\mathrm{A50},h,i,s}}$ and the $\mathrm{L}_{\mathrm{A50},h,i,s}$ 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 $\mathrm{L}_{\mathrm{A50},h,i,s}$ values less representative of the sound environment at a given location, (ii) $\overline{\mathrm{L}_{\mathrm{A50},h,i,s}}$ 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 $DANP_{i,s}$ 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 P_3 and the quietest one P_9. This variety of sound environments can also be illustrated by the fact that the $\overline{\mathrm{L}_{\mathrm{A50},h,i,s}}$ values at the loudest points at night are higher than that the $\overline{\mathrm{L}_{\mathrm{A50},h,i,s}}$ 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.

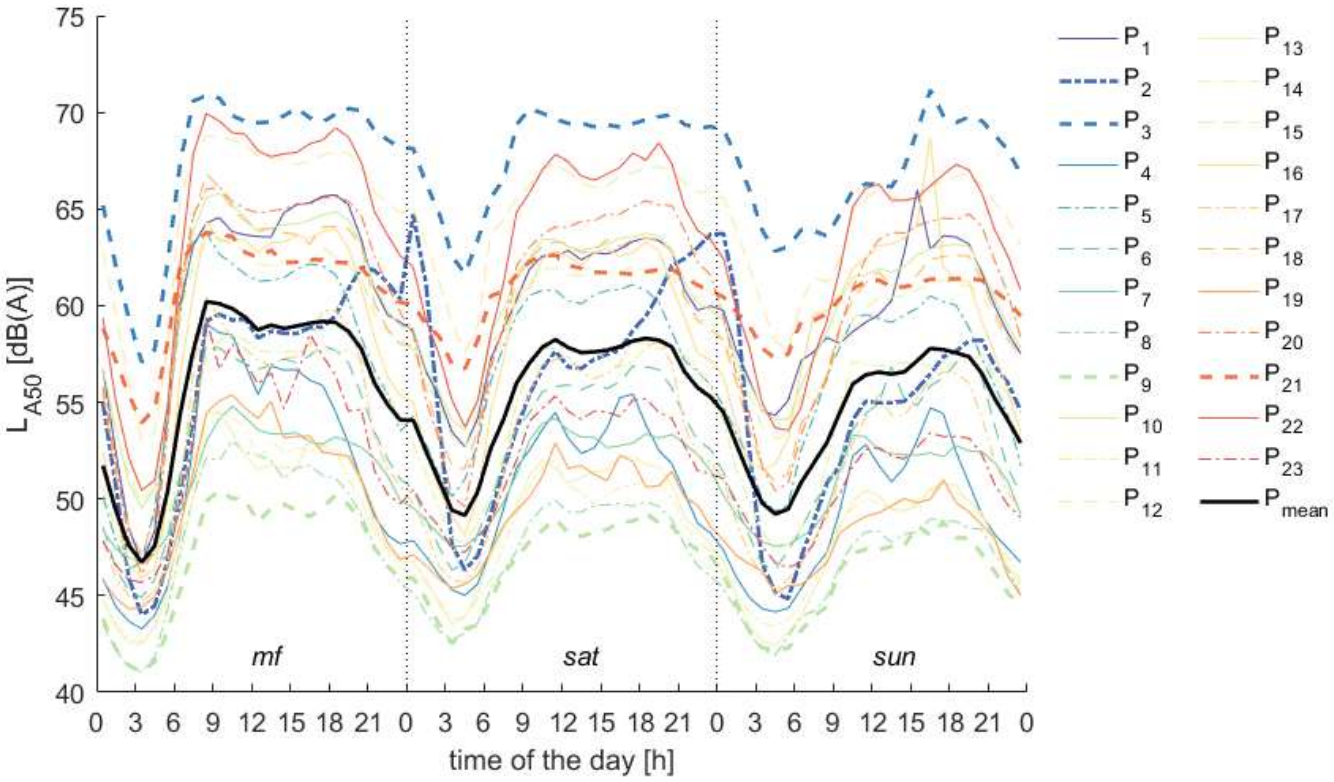

Figure 2. **L_{A50} Daily Average Noise Patterns at the 23 stations. The points discussed in the text are in bold.**

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 $\overline{L_{A50,h,i,s}}$ values 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 $\overline{r}$ (average of the 23 x 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 sound level time series at certain points. In particular, the sound level evolution at P_2 is poorly correlated to the other ones, showing high $\overline{L_{A50,h,i,s=P_2}}$ values in the *sat* and *sun* nights ($\overline{r_{sat}}$ = 0.58 and $\overline{r_{sun}}$ = 0.69). These high levels are explained by the particular location of P_2, 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. P_3 also shows low $\overline{L_{A50,h,i=sun,s=P_3}}$ correlations with the other points of the dataset ($\overline{r_{sun}}$ = 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.

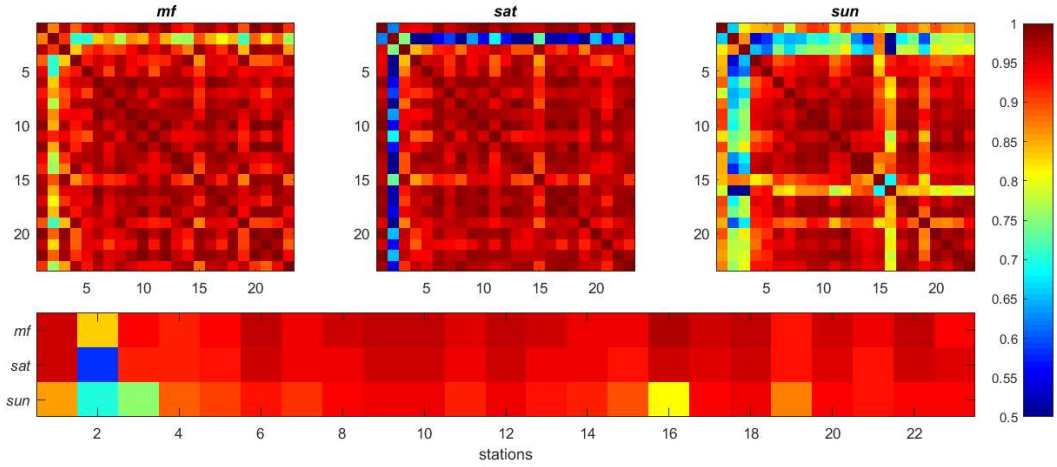

Figure 3. Correlations between the DANP at the 23 stations. Up: correlation matrices between each couple of stations for *mf*, *sat* and *sun*. Down: average correlation between each station and the 22 others .

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 $DANP_{i,s1} = a_{i,s1,s2} DANP_{i,s2} + b_{i,s1,s2}$, with s_1 and s_2 two monitoring stations. The interest of such modelling is to stretch the DANP through the $a_{i,s1,s2}$ parameter, and shift it through the $b_{i,s1,s2}$ parameter, and thus estimate $DANP_{s1}$ based on $DANP_{s2}$. 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*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 $DANP_{i,s1}$ and $DANP_{i,s2}$ from

the database. The domain of {a, b} values, represented in Figure 4 by dotted red lines, can be restrained to the following domain:

$$\begin{cases} b = 80 - 47.5\mathrm{a}, \\ b = 30 - 62.5\mathrm{a}, \\ 0.1 < \mathrm{a} < 4. \end{cases}$$

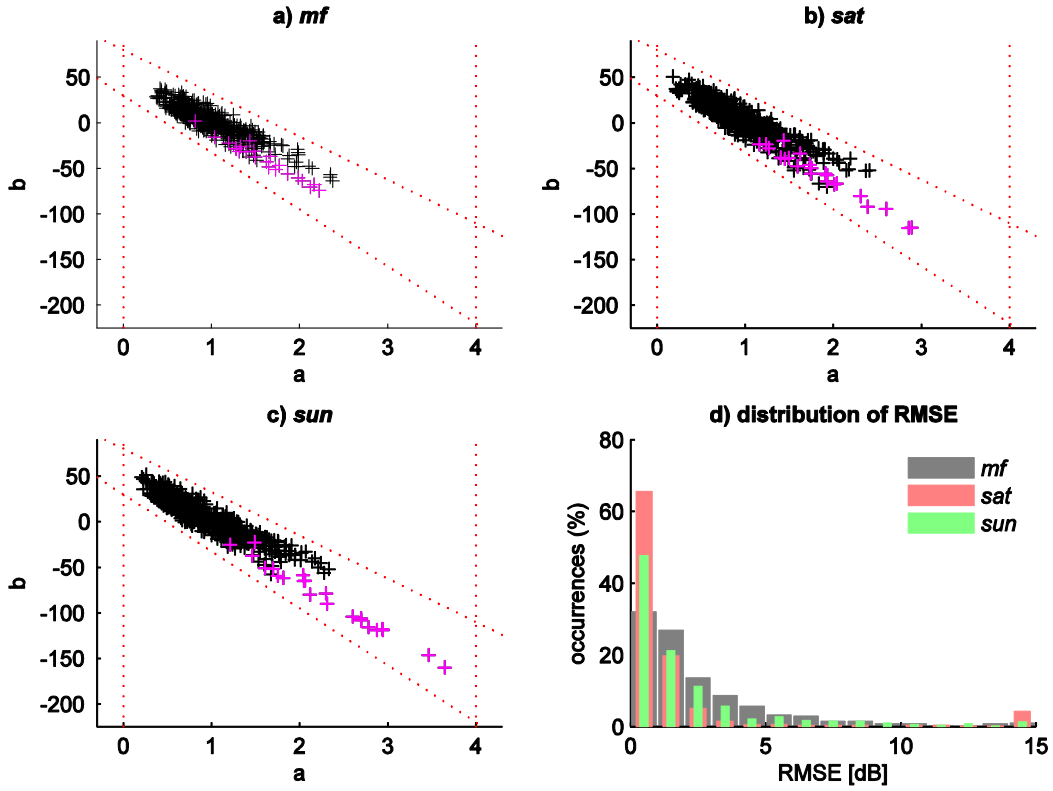

Figure 4. Domain of {a, b} values obtained when linking the DANP for all 506 possible combinations of two long-term monitoring stations through linear regression (crosses in magenta correspond to point P_{21}).

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 P_2, 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 $DANP_{i,s}$ based on samplings of a few $L_{A50,s}$ values at *s*, 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 $\delta_{i1,i2}$ are determined, with i_1 and i_2 = {*mf*, *sat*, *sun*}, each of the 9 matrices being of size 24 x 24, and each of its elements $\delta_{i1,i2}(h_1, h_2)$ containing the estimated delta value $\delta_{i1,i2}(h_1, h_2) = L_{A50,h1,i1} - L_{A50,h2,i2}$.

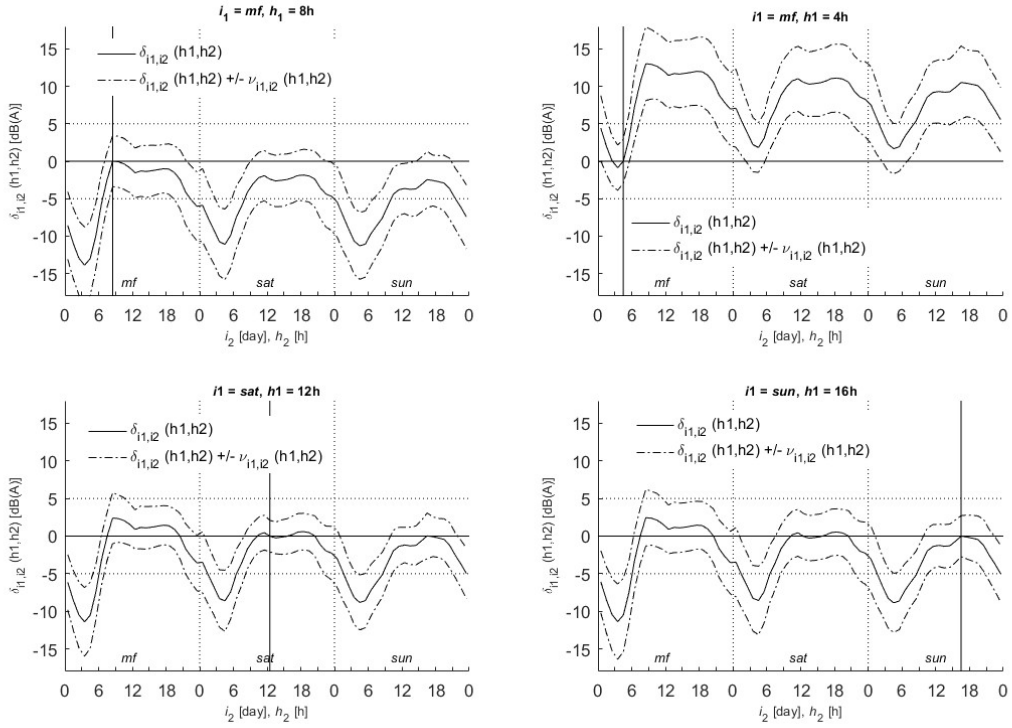

Figure 5. $\delta_{i1,i2}(h_1, h_2)$ and $\upsilon_{i1,i2}(h_1, h_2)$ values.

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 $\delta_{i1,i2}$ is associated with a matrix $\upsilon_{i1,i2}$ that gathers the standard deviation associated to the delta value. For example, $\delta_{mf,mf}$(10h,11h) = -0.3 dB and $\upsilon_{mf,mf}$(10h,11h) = 3.2 dB, whereas $\delta_{mf,sun}$(10h,05h) = -11.2 dB and $\upsilon_{mf,sun}$(10h,05h) = 4.1 dB. In practice, these $\delta_{i1,i2}$ and $\upsilon_{i1,i2}$ 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 ($L_{A50,h1,i1}$ - $L_{A50,h2,i2}$) values per month if i_1 and i_2 = *sat* or *sun* (supposing there is exactly four weeks

in the month), 16 * 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 * 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 $\delta_{i1,i2}(h_1,h_2)$ profiles. The profiles follow unsurprisingly the DANP shapes, with δ values ranging between -15 and +15 dB according to the $\{h_1, h_2\}$ couples. The diagonal of the $\delta_{i1,i2}$ matrix is null, but the diagonal of the $\upsilon_{i1,i2}$ matrix is not: for instance, $\upsilon_{mf,mf}$ (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 υ 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.

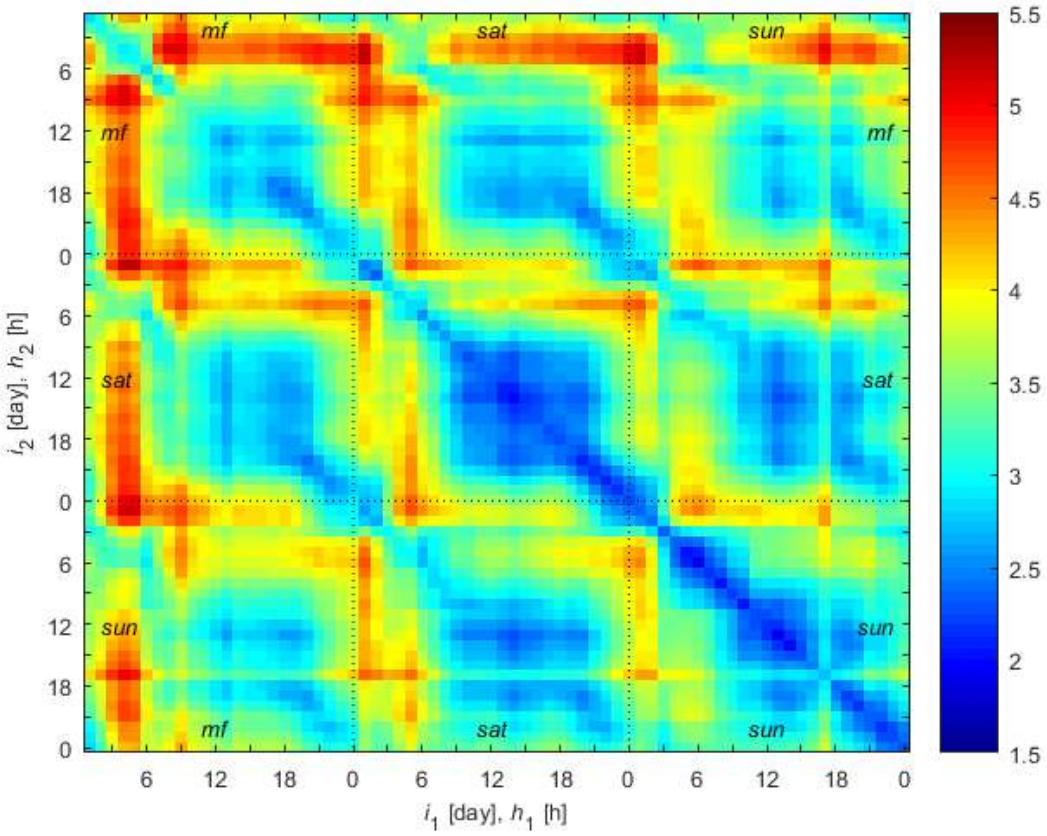

Figure 6. Matrix of $\upsilon_{i1,i2}(h_1, h_2)$ values.

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 stations for each 1h-time period for *mf*, *sat*, and *sun* (72 distributions per point). Generalized Extreme Value (GEV) distributions and normal distributions were compared for estimating the L_{A50} distributions. GEV distributions encompass the normal distributions but allow asymmetrical distributions, and are described by the following formula:

$$y = f(x|\kappa,\mu,\sigma) = \frac{1}{\sigma}\exp\left(-\left(1+\kappa\frac{x-\mu}{\sigma}\right)^{\frac{-1}{\kappa}}\right)\left(1+\kappa\frac{x-\mu}{\sigma}\right)^{-1-\frac{1}{\kappa}}, \quad (1)$$

where μ is the location parameter, σ is the scale parameter, and κ 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 fuction 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 P_1, 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 week-end 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 parameter. They will serve for estimating sound levels, especially in the model M_2 presented in section 4.3.

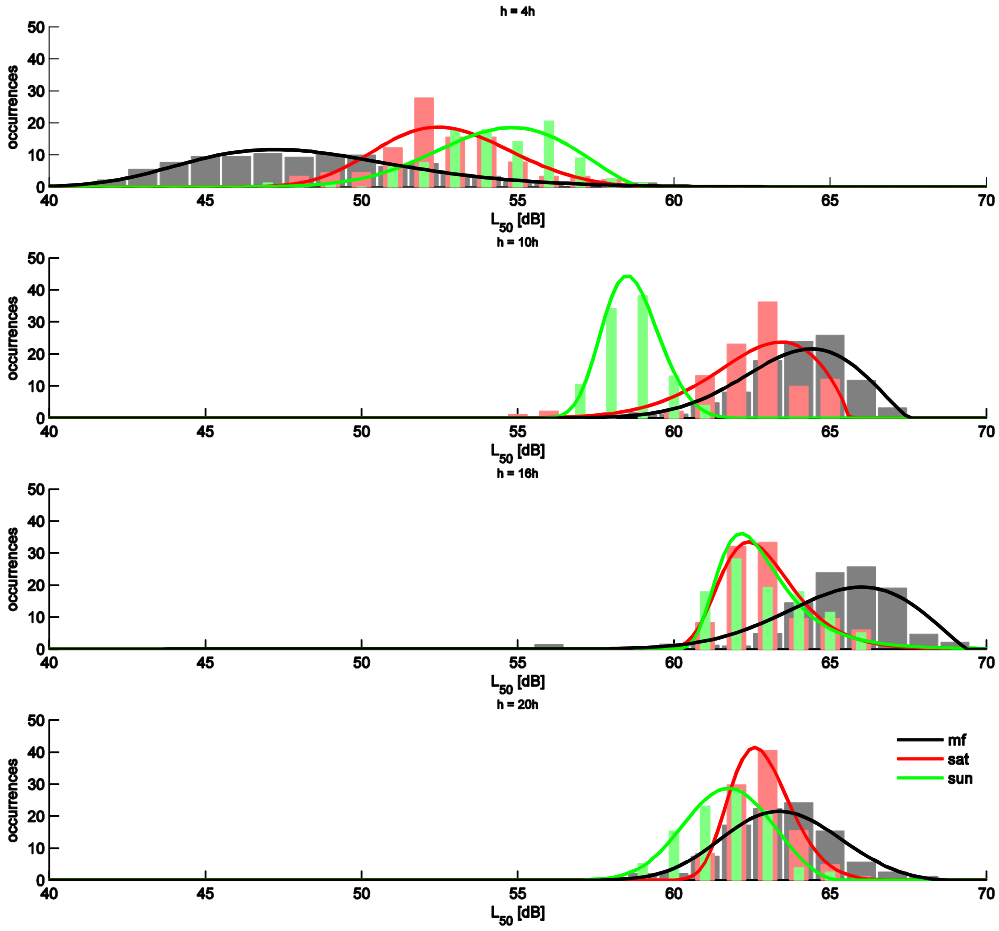

Figure 7. **Generalized Extreme Value distributions for *i* = {*mf*, *sat*, *sun*} at P_1 for the time periods *h* = 4h, *h* = 10h, *h* = 16h, and *h* = 20h.**

4. Modelling frameworks

Two modelling frameworks are proposed and compared, which rely on the information acquired from the long-term stations that are stored in the database, in order to estimate sound level patterns and probability density functions (pdf) at a new location with a limited amount of measures. The approach

consists of considering the sound level time series to be unknown at one monitoring station *S* among the 23 stations, and of using the data of the 22 remaining stations. One then tries to reconstruct the $DANP_{i,S}$ at *S*, knowing its actual reference value.

4.1. Simulated measurements

K virtual measurements are performed at *S*, which form a set of *K* $L_{A50,h(k),i(k),S}$ values, with *k* = {1; ...; K}, where h(k) is the time-of-the-day of the measure k. H(K) = {h(1)…∩…h(k)…∩…h(K)} is the set of the K time-of-the-day values. The duration of each measurement, which can be much shorter than 1h, is not a parameter of this study. Instead, each $L_{A50,1h}$ value is associated to a standard deviation σ_k. This standard deviation is the sum of the deviations due to the measurement protocol, the device, and the measurement duration. Information about the deviations due to measurement durations and to devices can be found in [26], and [11][18], respectively. In practice, the simulated measures are performed by taking randomly *K* measures among the 8-months database, and by adding a random value according to the standard deviation σ_k, which is a parameter of the study, to the known actual value. Measurement moments 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), $p_k(L_{50})$, which follows a normal distribution centred around $L_{A50,h(k),i(k),S}$ with a standard deviation of $\sigma_{k,h,i,S} = \sqrt{\left(\sigma_k + \sigma_{h,i,S}\right)^2}$.

4.2. Presentation of the Model M_1

The model M_1 relies both on the K $L_{A50,h(k),i(k),S}$ measurements at S and the delta matrices $\delta_{i1,i2}(h_1, h_2)$, in order to estimate the $DANP_{M1,i}$. First, each of the K $L_{A50,h(k),i(k),S}$ estimated values is extended to 72 probability density functions $p_{M1,k,h,i}(\mathrm{L_{A50}})$, with *h* = {1; …; 24} and *i* = {*mf*, *sat*, *sun*}, centred at $L_{A50,h(k),i(k),S} + \delta_{i,i(k)}(h, h(k))$, and with the standard deviation $\sigma_{k,h,i} = \sqrt{{\sigma_k}^2 + \upsilon_{i,i(k)}(h, h(k))^2}$. Then, at each 1h-time frame {*h, i*}, the K $p_{M1,k,h,i}(\mathrm{L_{A50}})$, are averaged to constitute the $p_{M1,h,i}(\mathrm{L_{A50}})$. The estimated $DANP_{M1,h,i}$ value is finally the centre of gravity of the pdf $p_{M1,h,i}(\mathrm{L_{A50}})$. 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 $DANP_{M1,i}$ that is more accurate at the time frames with many measurements, since the $\sigma_{i1,i2}(h1, h2)$ value is minimum when *h*1 = *h*2.

4.3. Presentation of the Model M_2

The Model M_2 relies both on the K $L_{A50,h(k),i(k),S}$ measurements at S and the $DANP_s$ of the stations *s* stored in the database to estimate the $DANP_{M2,i}$. First, The K $L_{A50,h(k),i(k),S}$ 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 $L_{A50,h(k),i(k),S}$ values and the corresponding $L_{A50,h(k),i(k),s}$ values at s :

$$w_{s,S} = \sqrt{\frac{K}{\sum_k \left(\mathrm{L}_{\mathrm{A50},h(k),i(k),S} - \mathrm{L}_{\mathrm{A50},h(k),i(k),s}\right)^2}}.$$

These similarities then serve to estimate $DANP_{M2,i}$ as a linear combination of the $DANP_s$ values, weighted by the $w_{s,S}$ values, such as $\mathrm{DANP}_{w,i} = \frac{\sum_s w_{s,S}\mathrm{DANP}_{s,i}}{\sum_s w_{s,S}}$. Probability density functions $p_{w,h,i}(\mathrm{L_{A50}})$ 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 $L_{A50,h(k),i(k),S}$ are determined. This linear regression aims to correct for the difference between the DANP at S and the $DANP_s$ 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 $DANP_{w,i}$ (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 $p_{w,h,i}(\mathrm{L_{A50}})$ probability density functions, to constitute the new pdf, $p_{r,h,i}(\mathrm{L_{A50}})$, which hopefully better reflects the sound level at S. Finally, the $p_{M_2,h,i}(\mathrm{L_{A50}})$ is at each {*h*, *i*} the average between the $p_{r,h,i}(\mathrm{L_{A50}})$ and the measurements $p_{k,h(k)=h,i(k)=i}(\mathrm{L_{A50}})$, which are achieved at {*h*, *i*}.

Thus, the model M_2 gives by construction an important weight to measurements at the time frames where numerous measurements are collected, since the measures only impact in M_2 estimates at their precise time frame, contrarily to the model M_1. 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

Case 1: σ_k = 1 dB

In Case 1, measurements are achieved at P_1 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 σ_k = 1 dB. The pdf functions and the DANP estimated with M_1 and M_2 are illustrated in Figure 8, for *K* = 5 measurements (left) and *K* = 20 measurements (right), along with the real DANP value at P_1 $DANP_{real}$, and the $DANP_{zone}$, which is the average of the 22 DANP.

First, this very low standard deviation of σ_k = 1 dB does not guarantee that measurements estimate the $\overline{L_{A50}}$ 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 $\overline{L_{A50}}$ 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 M_1 and M_2 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 P_1 than in the average of the other monitoring stations by about 5 dB. Sound levels are however slightly under-estimated, 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 M_2 seems to improve the noise level estimation during the night period, perhaps because the measurements help estimating the sound levels in M_2 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 M_2 better describes the high sound levels at P_1 during Sundays around 12:00, which are a particularity of the sound level evolution at P_1. This can be explained by the fact that sound level values at their time-period are specifically taken into account by the model.

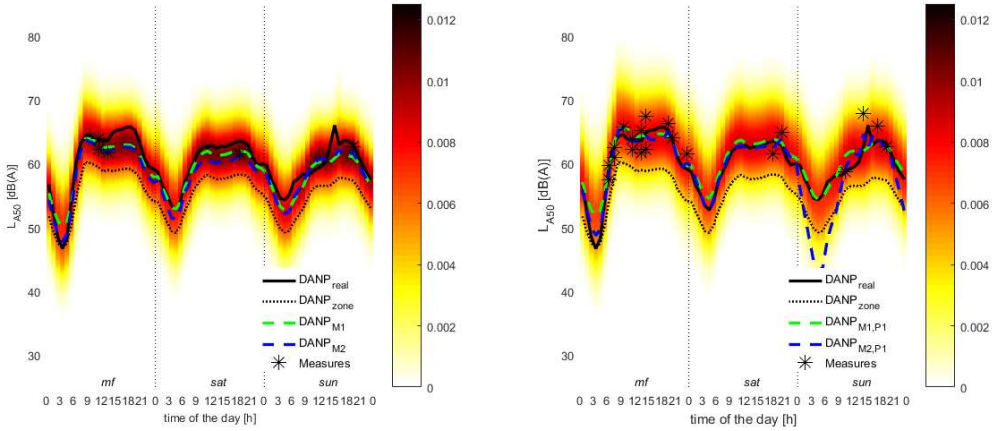

Figure 8. $DANP_{M1}$ and $DANP_{M2}$ estimates at P_1, with σ_k = 1 dB. Left: K = 5 measurements. Right: K = 20 measurements. The colorbar represents the probability density function.

Case 2: σ_k = 10 dB

In Case 2, measurements are obtained at P_1 with a standard deviation of σ_k = 10 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 dB. 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 P_1 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 σ_k = 10 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.

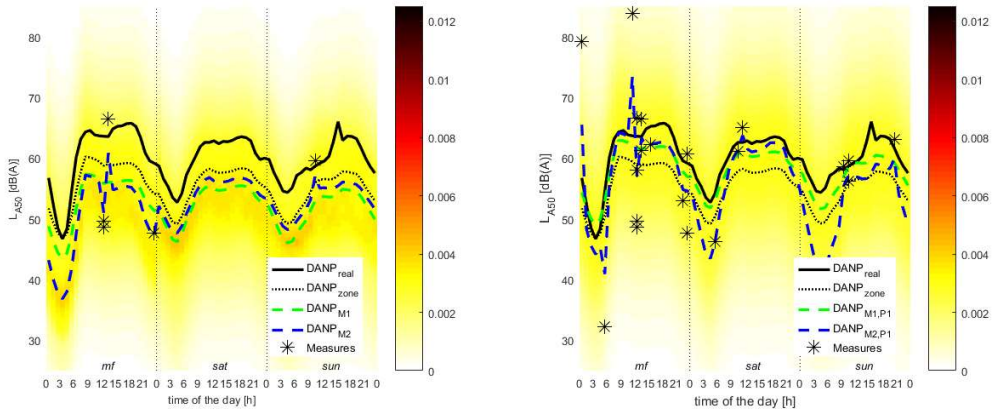

Figure 9. **$DANP_{M1}$ and $DANP_{M2}$ estimations at P_1, with σ_k = 10 dB. Left: K = 5 measurements. Right: K = 20 measurements.**

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 $WANP_s$ 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 $\overline{L_{A50,h,\iota,s}}$ values that form the estimated and the actual WANP ;
- the $\delta_{\overline{WANP}}$, which is difference between the estimated and the actual $\overline{WANP}$, which are the arithmetical averages of the 168 $\overline{L_{A50,h,\iota,s}}$ 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 σ_k = 1, 5 and 10 dB, according to the number of measurements.

In the case of very accurate individual measurements, when σ_k = 1 dB, a very low number of measures guarantees accurate WANP estimates, with $RMSE_{WANP}$ and $\delta_{\overline{WANP}}$ 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 $\overline{\mathrm{WANP}}$ estimates, the WANP are reproduced with a better accuracy by the model M_1, especially when the standard deviation associated to the individual measurements σ_k increases, confirming the conclusions from the analysis presented in section 4.4. The lower incidence of modelling choices on the $\delta_{\overline{WANP}}$ is certainly due to the compensations between the errors committed over the L_{A50} values that constitute the WANP. In any case, the model M_1 should be privileged when no information on the quality of measurements is available. As it captures efficiently the temporal structure of sound levels, the model M_1 permits a very precise estimation of both the noise levels and their temporal evolution, even with inaccurate measurements. As an example, with the model M_1 and with σ_k = 10 dB, K = 10 measurements are sufficient to estimate $\overline{\mathrm{WANP}}$ with an error of about 3 dB ($\delta_{\overline{WANP}}$ from 1 to 5 dB according to the points), and 15 measurements are sufficient to estimate the WANP with a RMSE_{WANP} of 3 dB.

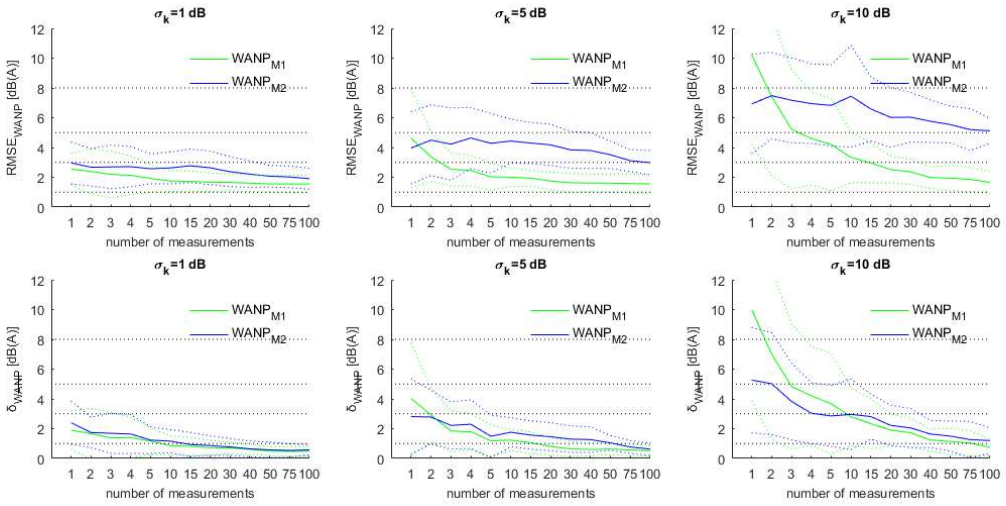

Figure 10. Comparison of models M_1 and M_2 according to the number of measurements. Left: σ_k = 1 dB. Centre: σ_k = 5 dB. Right: σ_k = 10 dB.

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 month. The time series of sound levels were recorded continuously with a 125ms-time resolution, from which $L_{A50,1h}$ 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_{A50} distributions, through the form of Generalized Extreme Values distributions. In addition, the average L_{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 $L_{A50,1h}$ 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 M_1 relies on the measurements along with the delta matrices, which help estimating DANP although not all time-periods are covered by measurements. The model M_2 relies instead on the DANP stored in the database, which are deformed thanks to the measurements through linear regressions, adding *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 M_1 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 *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.

7. Acknowledgements

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

8. References

[1] Directive 2002/49/EC of the European Parliament and of the Council of 25 June 2002, relating to the assessment and management of environmental noise. Official Journal of the European Communities. 2002.

[2] J. Morel, C. Marquis-Favre, L.A. Gille. Noise annoyance assessment of various urban road vehicle pass-by noises in isolation and combined with industrial noise: A laboratory study. Applied Acoustics **101** (2016) 47-57.

[3] A. Trollé, J. Terroir, C. Lavandier, C. Marquis-Favre, M. Lavandier. Impact of urban road traffic on sound unpleasantness: A comparison of traffic scenarios at crossroads. Applied Acoustics **94** (2015) 46–52.

[4] S. Bartels, F. Marki, U. Müller. The influence of acoustical and non-acoustical factors on short-term annoyance due to aircraft noise in the field — The COSMA study. Science of the Total Environment **538** (2015) 834-843.

[5] J.M. Wunderli, R. Pieren, M. Habermacher, D. Vienneau, C. Cajochen, N. Probst-Hensch, M. Röösli, M. Brink. Intermittency ratio: A metric reflecting short-term temporal variations of transportation noise exposure. Journal of Exposure Science and Environmental Epidemiology **26(6)** (2015) 1–11.

[6] A. Can, L. Leclercq, J. Lelong, J. Defrance. Capturing urban traffic noise dynamics through relevant descriptors. Applied Acoustics **69(12)** 2008 1270-1280.

[7] European Commission Working Group Assessment of Exposure to Noise (WG-AEN) - Position Paper - Final Draft - Good Practice Guide for Strategic Noise Mapping and the Production of Associated Data on Noise Exposure - Version 2 - 13th January 2006.

[8] S. Kephalopoulos, M. Paviotti, F. Anfosso-Lédée, D. Van Maercke, S. Shilton, N. Jones. Advances in the development of common noise assessment methods in Europe: The CNOSSOS-EU framework for strategic environmental noise mapping. Science of the Total Environment **482-483** (2014) 400-410.

[9] BruitParif, Site rumeur. http://rumeur.bruitparif.fr/.

[10] A. Can, L. Dekoninck, D. Botteldooren. Measurement network for urban noise assessment: Comparison of mobile measurements and spatial interpolation approaches. Applied Acoustics **83** (2014) 32-39.

[11] T. Van Renterghem, P. Thomas, F. Dominguez, S. Dauwe, A. Touhafi, B. Dhoedt, D. Botteldooren. On the ability of consumer electronics microphones for environmental noise monitoring. Journal of Environmental Monitoring **13** (2011) 544–552.

[12] E. D'Hondt, M. Stevens, A. Jacobs. Participatory noise mapping works! An evaluation of participatory sensing as an alternative to standard techniques for environmental monitoring. Pervasive and Mobile Computing **9(5)** (2013) 681-694.

[13] G. Guillaume, A. Can, G. Petit, N. Fortin, S. Palominos, B. Gauvreau, E. Bocher, J. Picaut. Noise mapping based on participative Measurements. Noise Mapping De Gruyter open journal **3** (2016) 140-156.

[14] M. Becker, S. Caminiti, D. Fiorella, L. Francis, P. Gravino, M. Haklay, A. Hotho, V. Loreto, J. Mueller, F. Ricchiuti, V.D.P. Servedio, A. Sîrbu, F. Tria. Awareness and Learning in Participatory Noise Sensing. PLoS ONE **8(12)** (2013) DOI: 10.1371/journal.pone.0081638.

[15] E. Kanjo. NoiseSPY: A Real-Time Mobile Phone Platform for Urban Noise. Monitoring and Mapping. Mobile Networks and Applications **15(4)** (2010) 562-574.

[16] C.A. Kardous, P.B. Shaw. Evaluation of smartphone sound measurement applications. Journal of the Acoustical Society of America **135(4)** (2014) 186-192.

[17] E. Murphy, E. King. Testing the accuracy of smartphones and sound level meter applications for measuring environmental noise. Applied Acoustics **106** (2016) 16–22.

[18] R.K. Rana, C.T. Chou, N. Bulusu, S. Kanhere, W. Hu. Ear-Phone: A context-aware noise mapping using smart phones. Pervasive Mobile Computing **17(A)** (2015) 1-22.

[19] A. Can, G. Guillaume, J. Picaut. Cross-calibration of participatory sensor networks for environmental noise mapping. **110** (2016) 99-109.

[20] S. Hachem, V. Mallet, R. Ventura, A. Pathak, V. Issarny, P.G. Raverdy, R. Bhatia. Monitoring noise pollution using the urban civics middleware. 2015 IEEE First international Conference on Big Data Computing Service and Applications.

[21] W. Wei, T. Van Renterghem, B. De Coensel, D. Botteldooren. Dynamic noise mapping: A map-based interpolation between noise measurements with high temporal resolution. Applied Acoustics **101** (2016) 127–140.

[22] A. Can, G. Guillaume, B. Gauvreau. Noise indicators to diagnose urban sound environments at multiple spatial scales. Acta Acustica united with Acustica **101** (2015) 964-974.

[23] A.J. Torija, D.P. Ruiz, A. Ramos-Ridao. Required stabilization time, short-term variability and impulsiveness of the sound pressure level to characterize the temporal composition of urban soundscapes. Applied Acoustics **72(3)** (2011) 89–99.

[24] C.P. Gajardo, J.M. Barrigon Morillas, G.R. Gozalo, R. Vilchez-Gomez. Can weekly noise levels of urban road traffic, as predominant noise source, estimate annual ones? The Journal of the Acoustical Society of America **140(5)** (2016) 3702-3709.

[25] A. Can, T. Van Renterghem, M. Rademaker, S. Dauwe, P. Thomas, B. De Baets, D. Botteldooren. Sampling approaches to predict urban street noise levels using fixed and temporary microphones. Journal of Environmental Monitoring **13** (2011) 2710-2719.

[26] L. Brocolini, C. Lavandier, M. Quoy, C. Ribeiro. Measurements of acoustic environments for urban soundscapes: choice of homogeneous periods, optimization of durations, and selection of indicators. Journal of the Acoustical Society of America **134 (1)** (2013) 813-821.

[27] P. Aumond, A. Can, C. Lavandier, B. De Coensel, D. Botteldooren, C. Ribeiro. Modeling soundscape pleasantness using perceptive assessments and acoustic measurements along paths in urban context. Acta Acustica united with Acustica (submitted for publication).

[28] P. Ricciardi, P. Delaitre, C. Lavandier, F. Torchia, and P. Aumond, "Sound quality indicators for urban places in Paris cross-validated by Milan data," Journal of the Acoustical Society of America **138(4)** 2015 2337–2348.

[29] H. Kuchenhoff, M. Thamerus, "Extreme value analysis of Munich air pollution data," Sonderforschungsbereich **386(4)** (1995).

[30] L.D. Martins et al., "Extreme value analysis of air pollution data and their comparison between two large urban regions of South America," Weather and Climate Extremes (2017), https://doi.org/10.1016/j.wace.2017.10.00.

[31] Z. Rulfová, A. Buishand, M. Roth, J. Kyselý, "A two-component generalized extreme value distribution for precipitation frequency analysis," Journal of Hydrology (534) (2016) 659-668.

[32] D. Geraghty, M. O'Mahony. Investigating the temporal variability of noise in an urban environment. International. Journal of Sustainable Built Environment **5** (2016) 34-45.

9. List of abbreviations and variables

$DANP_{i,s}$: Daily Average Mean Pattern at location s, which consists on 24 $\overline{L_{50,h,i,s}}$ values

GEV : Generalized Extreme Value distribution

h : time-of-the-day, with *h* = {0; 1; … ; 23}

H(K) : set of the K time-of-the-day values

i = {*mf*, *sat*,*sun*} : days of the week, where "*mf*" stands for one of the five days from Monday to Friday, "*sat*" stands for Saturday, and "*sun*" stands for Sunday.

$L_{A50,1h}$: median of the $L_{Aeq,1s}$ values among a 1h period

$\overline{L_{A50,h,i,s}}$: average of the $L_{A50,1h}$ values collected at *s* during the time-of-the-day *h*, and the day-of-the-week *i*.

M_m : model M_m, with m = {1,2}

pdf: probability density function

$P_{L_{A50,1h}}$: Probability density function associated to an estimated $L_{A50,1h}$ value

$RMSE_{DANP}$: difference between the estimated and the actual DANP

s = {1; …; 23}: 23 monitoring stations.

$\delta_{i1,i2}$: matrix of size 24*24 that contains the estimated delta value $\delta_{i1,i2}\ (h1, h2) = L_{50,h1,i1} - L_{50,h2,i2}$

$\delta_{\overline{WANP}}$: difference between the estimated and the actual $\overline{\mathrm{WANP}}$, which are the arithmetical averages of the 168 $\overline{\mathrm{L_{A50,}}_{h,i,s}}$ values that form the estimated and the actual WANP.

$\kappa_{h,i,s}$: GEV distribution shape parameter at station *s*, day-of-the-week *i*, and time-of-the-day *h*.

$\mu_{h,i,s}$: GEV distribution location parameter at station *s*, day-of-the-week *i*, and time-of-the-day *h*.

$\upsilon_{i1,i2}$: matrix of size 24*24 that contains the standard deviation of the differences $\delta_{i1,i2}\ (h1, h2)$

$\sigma_{h,i,s}$: GEV distribution scale parameter at station *s*, day-of-the-week *i*, and time-of-the-day *h*.

σ_k : standard deviation associated to measurements

$w_{s,S}$: similarity (or weight) estimated between two stations s and S

WANP: concatenation of 5 $DANP_{mf}$, one $DANP_{sat}$ and one $DANP_{sun}$, thus representing the one-week average noise pattern.