Optimization of Dataset Generation for a Multilinear Regressive Road Traffic Noise Model

DOMENICO ROSSI, AURORA MASCOLO, CLAUDIO GUARNACCIA
Department of Civil Engineering,
University of Salerno,
Via Giovanni Paolo II, 132 – 84084 Fisciano (SA),
ITALY

Abstract: - According to the European Environmental Agency, road traffic noise is one of the worst and most prevalent kinds of environmental pollutants, which causes health problems to a constantly increasing number of people in urban areas throughout Europe. It has been proved that prolonged exposure to sound levels exceeding 55 dBA is harmful and causes severe problems like sleep disturbances, tiredness, lack of concentration, high blood pressure and, in the worst case, sudden death. A precise and constant evaluation of sound level in inhabited areas is therefore desired (and in some cases compelled by laws), but collection of actual noise data is not easy and sometimes not possible. For this reason, Road Traffic Noise (RTN) models are very handy: one can (more or less precisely) estimate the noise emitted in a certain area having certain road traffic characteristics. The application of RTN models, anyway, also has problems. First of all, an RTN model has to be built and calibrated by using real collected noise data. Moreover, when trying to apply an RTN model on road traffic situations that are far away from the site of collection, the models generally fail. To overcome such problems, in this contribution, a road traffic dataset has been computed by randomly generating values of traffic variables like the number of vehicles per unit of time, their velocities, and their distance from the receiver. Then, by applying a multiregressive function on the dataset, the obtained coefficients have been used to calibrate and validate the presented model. The three steps (generation of the dataset, calibration of the model, and validation on a real dataset) are detailly investigated.

Key-Words: - Road Traffic Noise Model, Multilinear Regression, Computed Calibration Dataset, Sensitivity Analysis, Outlier Analysis, Data Trimming

Received: March 21, 2023. Revised: August 23, 2023. Accepted: October 15, 2023. Published: November 9, 2023.

1 Introduction

Road Traffic noise is one of the most intrusive kinds of noise in urban contexts. European Environment Agency has estimated that a large part of the population is constantly exposed to noise levels exceeding the safety threshold (55 dBA). If prolonged, such exposure can lead to several health issues like annoyance, sleep disturbances, high blood pressure, and even sudden death, [1], [2], [3], [4]. In urban contexts noise is mainly, but not exclusively, constituted by traffic, [1], [5], which has been growing over the years. Road traffic noise – i.e. the one generated by passing vehicles - is not the only one contributing to the high noise level in urban contexts. Sources other than the cars, for example, are railway noise, which is also recognized to be detrimental to human health, [6], noise coming from port areas, [7], and even motor race events, [8], when the circuit is not adequately far away from the urban center. Back to road traffic noise, to determine the noise level of a specific area, a campaign of measurements with specific sound level meters must be organized, but such an approach is not always applicable. For different reasons, in fact, real measurements can be not available, and alternative approaches have to be followed. Typically, when the assessment of the noise level in a specific area is not possible because of a lack of real measured data, Road Traffic Noise (RTN) models can supply by simulating noise levels according to independent parameters, which usually are the number of passing vehicles, their speed and the distance between each noise emitting vehicle and the receiver. Some models also take into account parameters like climatic conditions (which can modify the noise propagation), structure of acoustic barriers, or specific road conformation (like the presence of roundabouts, [9], [10]). Several models for road traffic noise estimation are present, and they are generally framed into national laws and regulations. Some example are: the CoRTN model in United Kingdom, [11], the RLS90 model used in

Germany, [12], the NMPB model in France, [13], the ASJ model in Japan, [14]. Another model important to mention is the CNOSSOS model, which is recently born by the efforts of the European Community to create a model including several noise sources, [15]. Application of RTN Models, anyway, also have significant problems, since they need a proper calibration, which is, in turn, usually performed basing on real measured data. Moreover, such calibration process typically makes the model – when properly built – usable in the same context where the calibration data have been collected, but less efficient when applied on a different site. For this reason, road traffic noise models built in a country are adopted from the same country but not from others. To overcome such issues, the authors conceived an RTN model (based on a multilinear regression) which is calibrated on computed data instead of real collected ones. By a proper tuning of the parameters and hyperparameters of the functions generating the data it is in fact possible to obtain a dataset mimicking a wide range of real situations. The subsequent calibration, then, results in a model not anymore bounded to simulate traffic conditions in a single place. The generation of the computed set of data took idea from a previously published work, [16], and further improved. In this contributions authors investigated how the manipulation of data in terms of dataset size, different sets of data and time of execution can modify the final efficiency of the model in simulating road traffic noise levels. IN detail, the model has been validated using measured data coming from a work of Université Gustave Eiffel and *Unité Mixte de Recherche en Acoustique Environnementale* (UMRAE), Nantes, called "Long Term Monitoring Station" (LTMS). LTMS was an experimental campaign of collection of acoustic and meteorological data on a road of the city of Saint-Berthevin, made from 2002 to 2007, [17]. The final dataset is available for research purposes.

2 Material and Methods

2.1 Computation of the Dataset for Model Calibration

The computed datasets used for the implementation and testing of the here presented multi-regressive model have been generated on a DELL Pc (Intel® Xeon® CPU E3-1245 v5 @3.50 GHz with 16 Gb of RAM installed, 64bit) using Python, a free objected oriented programming language. Several Python packages have been used for the generation of the dataset: the more important were numpy, which is a numerical package for calculations, pandas, which is a package for the creation, organization, and filtering of datasets, and matplotlib. pyplot and seaborn, which are packages used for the plotting of data. The compiler chosen for running the Python code is Jupyter Notebook, which works with a Google browser interface and permits the organization of the script at isolated blocks so that the written code can be run after being sliced in pieces. The generation of the dataset proceeded per row, by filling each of the independent variables with a randomly extracted value within predetermined ranges. The exact sequence is reported in Figure 1, and it has been established as follows.

2.1.1 Generation of Independent Variables

1) *Determination of Q.* Flow, expressed as vehicles passing on the investigated road per time period, has been chosen to cover all the possible situations, spanning from a minimum of 10 vehicles/time to a maximum of 2000 vehicles /time, with a step of 10 vehicles /time.

2) *Random extraction of light vehicles velocity.* The velocity of light vehicles (common cars) is randomly extracted from a range spanning a minimum of 30 km/h to a maximum of 130 km/h, with a step size of 1 km/h. Please note that all the possible velocities between the minimum and the maximum range have the same probability to be extracted, and this characteristic applies in all the following points regarding the random extraction of values (2 to 7).

3) *Random extraction of medium vehicle velocity.* The velocity of medium vehicles is randomly extracted from a minimum of 30 to a maximum of 100 km/h, with a step size of 1 km/h. If the velocity of light vehicles falls within this range, such a value becomes the upper limit for the random extraction of the medium vehicles' velocity.

4) *Random extraction of heavy vehicles velocity.* The velocity of high vehicles is randomly extracted from a minimum of 30 to a maximum of 80 km/h, with a step size of 1 km/h. As for the medium vehicles, if the velocity of light vehicles falls within this range, such a value becomes the upper limit for the random extraction of the medium vehicles' velocity.

5) *Random extraction of medium vehicles percentage.* The percentage of medium vehicles over the total Q is randomly extracted from a minimum of 0% to a maximum of 20%, with a step size of 0.1%.

6) *Random extraction of heavy vehicles percentage*. The percentage of medium vehicles over the total Q is randomly extracted from a minimum of 0% to a maximum of 20% minus medium vehicles percentage, with a step size of 0.1%.
7) *Random extraction of distance.* The distance between noise-emitting cars and the receiver is randomly extracted from a minimum of 10 to a maximum of 100 m, with a step size of 1 m.
8) *Multiplication of the row number*. Steps from 2 to 7 are repeated n times to statistically enlarge the dataset.
9) *Simulation of the* $L_{eq,t}$. $L_{eq,t}$ values are simulated by using REMEL, [18], as a noise emission model and the formulation found in [19]. In detail: $L_{eq,t}$ values are calculated with a specific Noise Emission Model (NEM). The NEM used in this case study is the one presented in the work of [18], [19]. In detail, as the first step REMELs (the sound level at a referring distance) are calculated for light, medium, and heavy vehicles by using the following equations:

$$L_{0,L} = 31.13 \log_{10}(vL) + 12.777 \quad (1)$$
$$L_{0,M} = 18.765 \log_{10}(vM) + 43.697 \quad (2)$$
$$L_{0,M} = 18.765 \log_{10}(vM) + 43.697 \quad (3)$$

where L0 are the power levels at the source. Then, the actual sound power level of a single vehicle traveling at a specific velocity is propagated by equations 4, 5, and 6.

$$\mathrm{L_{WL}} = \mathrm{L_{0,L}} + 20 \log_{10} \mathrm{D_{ref}} + 11 \quad (4)$$
$$L_{WM} = L_{0,M} + 20 \log_{10} D_{ref} + 11 \quad (5)$$
$$L_{WH} = L_{0,H} + 20 \log_{10} D_{ref} + 11 \quad (6)$$

being D_{ref}, a reference distance of 15 m. In the third step, the Sound Exposure Levels are calculated with formulas 7, 8, and 9. SEL is the total sound emitted by the vehicles, as it was emitted in one single second.

$$SEL_L = 10 \log_{10} Q_L + L_{WL} - 20 \log_{10} d + 11 \quad (7)$$
$$SEL_M = 10 \log_{10} Q_L + L_{WM} - 20 \log_{10} d + 11 \quad (8)$$
$$SEL_H = 10 \log_{10} Q_H + L_{WH} - 20 \log_{10} d + 11 \quad (9)$$

Q_L, Q_M, and Q_H are respectively the number of light, medium, and heavy vehicles, and d is the distance as randomly generated in the phases of dataset construction (please refer to step A of this section). Finally, the total equivalent level ($L_{eq,t}$) is computed as follows (equation 9)

$$L_{eq,t} = 10 \log_{10}\left(\frac{1}{sec}\right) + 10 \log_{10}(10^{\frac{SEL_L}{10}} + 10^{\frac{SEL_M}{10}} + 10^{\frac{SEL_H}{10}}) \quad (10)$$

In equation (10), sec is the number of seconds used to evaluate the final equivalent sound level, which is the sound level at which a receiver is exposed at a certain distance d for a certain period of time. To be compliant with regulations and literature, a typical value of 3600 seconds (1 hour) has been chosen. Once the hourly equivalent level ($L_{eq,h}$ from now on) is computed, the multilinear regression can be applied by setting $L_{eq,h}$ as the dependent variable. Correlation between each variable and $L_{eq,h}$ is established and the coefficients are stored. Naming C1, C2, C3 and so on the regression coefficients, a final equation is used to validate the model on a real dataset:

$$L_{eq,h\ simulated} = C1vL + C2vM + C3vH + C4PM + C5PH + C6d + intercept \quad (11)$$

2.1.2 Variation of Dataset Size

Dataset size can be varied, according to the previously presented schematization, by varying n, which is the time the random extraction of the independent variables is performed, associated with each Q value. The authors, then, performed the generation of a single dataset by varying the hyperparameter n, chosen values of n are 1, 2, 5, 10, 30, 60. Since Q values are originally 200, the corresponding resulting dataset has a number of rows equal to 200, 400, 1000, 2000, 6000, and 12000. The generated dataset is then calibrated with the usual multilinear regression technique, obtaining calibration residuals, and then validated on the same validation dataset. Computation of the datasets requires a variable time, which increases at the increase of the n value. Analysis of such computing time has been performed with the in-built Jupyter Notebook %%time function, that gives back, after each block of compiled code, the wall time and the CPU time. CPU time is the time actually spent by the CPU in the execution of the code (for this reason it is sometimes referred to as "execution time"), whereas wall time is the real time elapsed between the code run and the visualization of the result. Such last parameter is greatly affected by the business of the CPU since it is slower when other processes are running.

2.1.3 Variation of Used Data (seed)

Another useful in-built function of the used packages (numpy and pandas) is the seed function, which permits to tracing of the random choice of values. When randomly placing the values of independent variables during dataset generation, in fact, it is useful to store a seed number that permits to recreate of the exact combination of

independent variables' values and makes the dataset precisely reproducible.

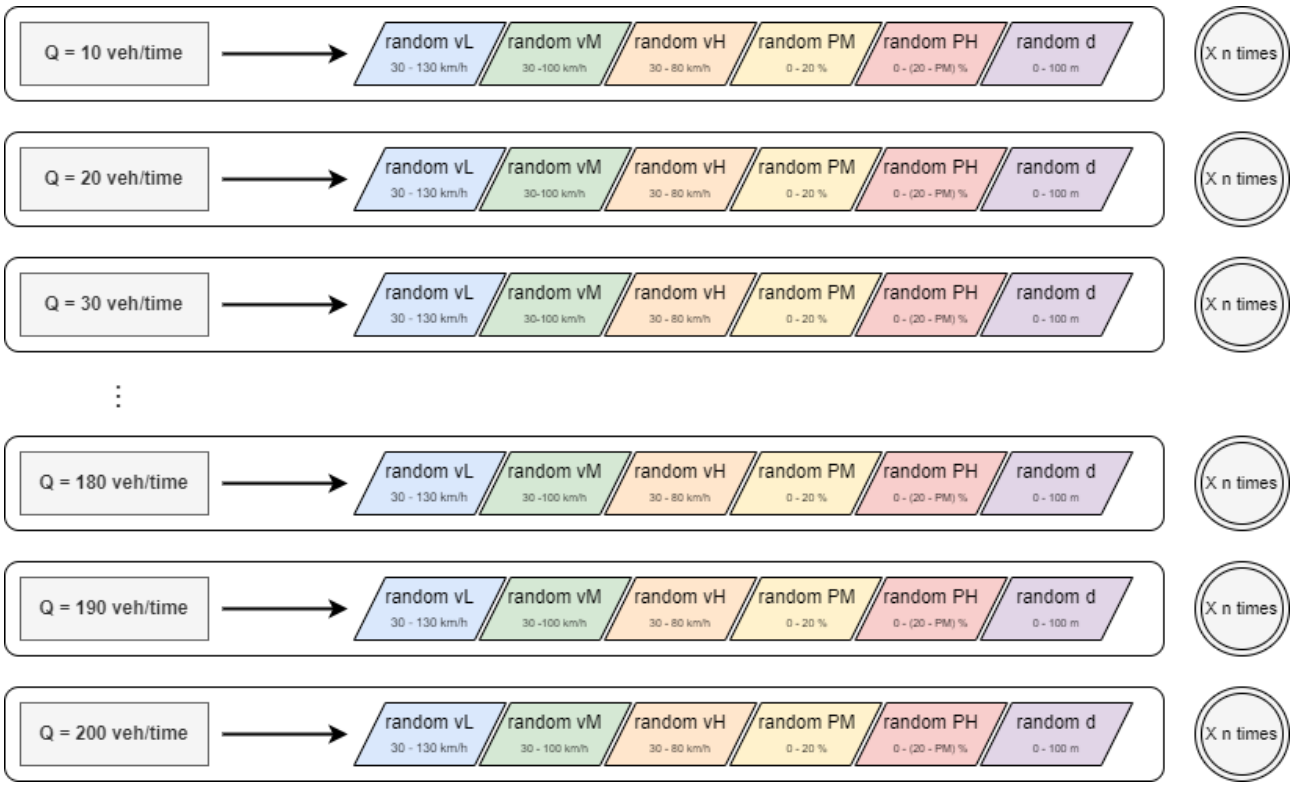

Fig. 1: Scheme of the generation of the dataset for model calibration. Each independent variable is randomly extracted within certain ranges and is associated with a fixed Q value spanning from 10 to 2000 vehicles per time period. The random extraction of some of the variables is constrained by the value of other variables.

When executing a function, the seed parameter is declared between the parameters of the function itself as a number; each time the function is run declaring the same seed number the results will be exactly the same. Such operation is very important to test different algorithms on the same data and compare the results, but it is also essential to fulfill the scientific purpose of reproducibility of an experiment. Every single dataset, then, has been generated 100 different times by using 100 different seeds (for the sake of practicality the seed number are from 1 to 100) to verify whether the variation of data (due to different used seeds) can be associated to the diverse final result of the model in terms of calibration and validation.

2.2 Calibration Step

Multilinear regression technique has been applied to the generated dataset to determine the proper coefficient and slope to be used to correlate the $L_{eq,t}$ values to the independent variables. When performing multilinear regression, a population of residuals is obtained, which is calculated by the difference between the real data and the calibrated one. When the model is properly calibrated and no biases are present, the population of residuals is normally shaped with an average value around 0 and a certain sigma value.

2.3 Statistical Analysis of the Generated Dataset

Generated datasets have been statistically investigated by means of several parameters. The mean and standard deviation of the populations of residuals have been investigated, together with the Shapiro-Wilk test for the normality of the distribution.

2.3.1 Outliers Trim and New Iteration of Multilinear Regression

The application of the multilinear regression to the dataset generates a series of calibration residuals, which are defined as the discordances between the values obtained after the calibration and the original values. The distribution of such residuals is indicative of the goodness of the calibration process. To improve the calibration process, the multilinear regression has been performed twice. First multilinear regression has been performed on the overall dataset, obtaining the first residual

population (step D). Such residual population has been investigated by searching the outliers and removing them. Outliers in the residuals population are the expression of the row of the original dataset where the calibration process got a value significantly different from the original one. By removing the outliers, then, the authors pruned the original dataset eliminating the rows in the multilinear regression that had more difficulty in fitting. A second multilinear regression, then, has been applied to the dataset diminished by the removal of the outliers, calculating a new residual population. The process of outliers removal was performed by trimming the data that fall outside the range defined in the following formula (12):

$$mean = iqr \times k_0 \quad (12)$$

where iqr is the interquartile range, i.e. the difference between the 75th and the 25th quantile of the residuals population and k_0 is a factor that has been iterated from a minimum of 0.1 to a maximum of 4.0.

2.4 Validation of the Model

These final simulated equivalent levels are compared with the real ones to assess the sensitivity of the model. The real dataset used for the validation of the model is the Long Term Monitoring Station (LTMS) dataset. LTMS dataset is a collection of environmental acoustics data collected by a system installed in the city of Saint-Berthevin (France) by the Université Gustave Eiffel and Unité Mixte de Recherche en Acoustique Environnementale (UMRAE), Nantes, over a period of ten years. The authors in [17], suggested that LTMS is originally composed of equivalent sound levels measured over 15 minutes, so data preprocessing has been performed to convert the data in 1 one-hour time period (the procedure described until now has been also visually explicated in Figure 2).

The results obtained are compared with the real $L_{eq,h}$ values by comparing the two data distributions in terms of mean, standard deviation, skewness, and kurtosis index. Then the Mean Absolute Error (MAE) and Mean Absolute Percentage Error (MAPE) values are used to assess the final sensitivity of the model itself.

3 Results and Discussion

3.1 Generation of the $

The generation of the dataset for the calibration of the multilinear regression model is a crucial point for the sensitivity of the model itself. The populations of the dataset variables have been investigated through statistical analysis, which is reported in Table 1. The numerosity of the sample, mean and standard deviation of the sample, skewness, and kurtosis index of the distribution and Shapiro-Wilkins test have been performed on the datasets built by varying the n number (i.e. by varying the sample numerosity) As visible, the statistical analysis shows no difference between the datasets, meaning that the procedure of the generation of the dataset, i.e. the random extraction of the values of independent variables and the constraints between the variables themselves (see section 2 “Material and methods” and also, [16]) is solid, is not affected by the numerosity of the sample and it does not introduce bias on the samples.

3.2 Calibration of the Model

The obtained datasets were processed with the multilinear regression function described in section three, in order to find the residual population and evaluate whether the dataset amount influenced the calibration phase by the residuals analysis. Table 2 collects the results of the application of the already used statistical parameters on the residual population of each of the generated datasets. In all cases, the mean of the population is around 0.0, but for the other parameters, some discrepancies are visible. The standard deviation of the population, for instance, decreases with increasing of the dataset size, suggesting that the generation of a bigger dataset improves the width of the residuals’ distribution. The other parameter with a similar behavior is the kurtosis index, which increases with the dataset size, from a minimum value of 0.504 up to a maximum tested value of 1.456 for the 60X dataset. This means that the residual population becomes “sharper” when the original dataset has more entries, and since the mean of the distributions is always 0.0, the distributions are more centered and fewer values fall on the tails region.

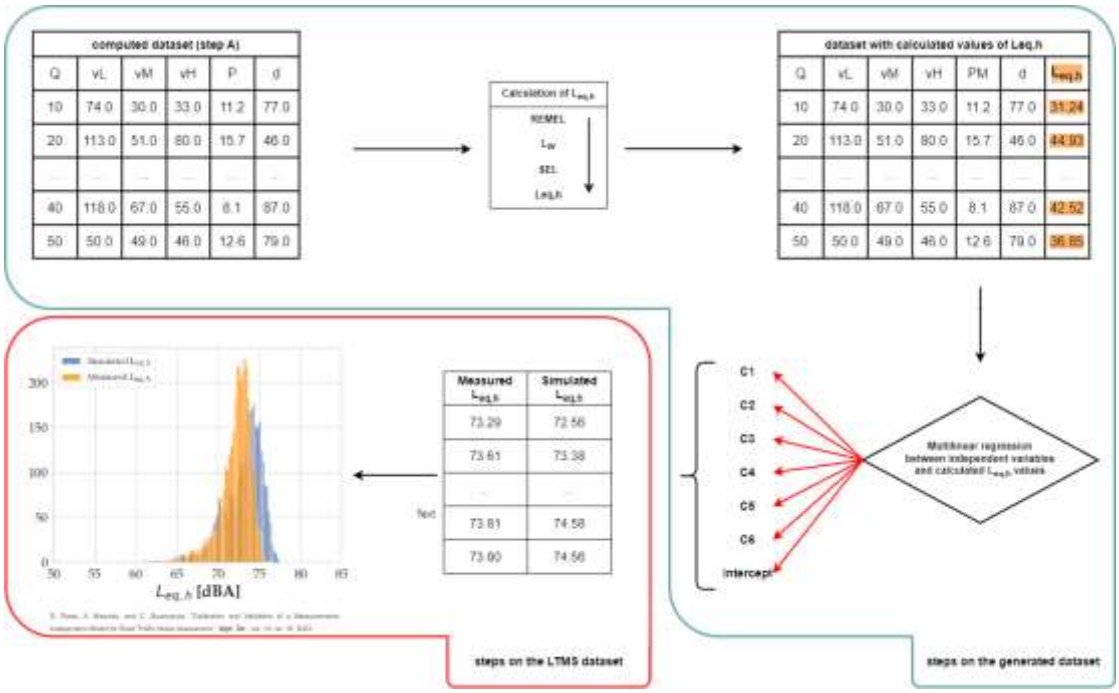

Fig. 2: schematic representation of the steps of $L_{eq,h}$ generation with the used NEM and the subsequent steps of calibration (by using multilinear regression technique) and validation of the model to the real data of LTMS.

Table 1. Statistical parameters of the variables of the dataset at different sizes of the dataset itself.

	1X	2X	5X	10X	30X	60X
Q						
Mean [veh/time]	1005	1005	1005	1005	1005	1005
Std [veh/time]	578.792	578.067	577.632	577.487	577.391	577.367
skewness	0.0	0.0	0.0	0.0	0.0	0.0
kurtosis	-1.2	-1.2	-1.2	-1.2	-1.2	-1.2
Shapiro	0.955	0.955	0.955	0.955	0.955	0.955
vL						
Mean [km/h]	82.875	83.432	83.314	82.971	82.341	82.036
Std [km/h]	26.669	30.269	30.049	30.114	30.122	30.226
skewness	-0.036	-0.064	-0.065	-0.054	-0.016	0.002
kurtosis	-1.227	-1.264	-1.188	-1.176	-1.179	-1.192
Shapiro	0.952	0.948	0.955	0.956	0.956	0.954
vM						
Mean [km/h]	54.13	54.608	530.814	53.876	53.343	53.285
Std [km/h]	17.941	18.559	19.195	19.305	193194	19.148
skewness	0.583	0.619	0.670	0.648	0.682	0.683
kurtosis	-0.608	-0.627	-0.620	-0.669	-0.610	-0.615
Shapiro	0.941	0.933	0.921	0.926	0.917	0.917
vH						
Mean [km/h]	50.08	50.338	50.380	49.856	49.811	49.880
Std [km/h]	15.725	15.885	15.802	15.621	15.55	15.570
skewness	0.514	0.485	0.490	0.528	0.537	0.542
kurtosis	-0.913	-0.953	-0.951	-0.881	-0.859	-0.855
Shapiro	0.925	0.927	0.927	0.926	0.927	0.927
P						
Mean [%]	15.092	15.030	14.952	14.955	14.932	15.007
Std [%]	3.996	4.143	4.475	4.470	4.526	4.435
skewness	-0.799	-0.782	-0.931	-0.980	-0.960	-0.971
kurtosis	-0.030	-0.169	0.033	0.195	0.080	0.160
Shapiro	0.927	0.921	0.899	0.897	0.895	0.897
d						
Mean [m]	52.98	54.243	54.674	55.301	55.658	55.013
Std [m]	26.320	26.108	26.012	26.246	26.301	26.241
skewness	0.124	0.002	0.027	-0.022	-0.034	-0.002
kurtosis	-1.131	-1.133	-1.154	-1.184	-1.20	-1.199
Shapiro	0.954	0.958	0.958	0.956	0.954	0.955

Table 2. Statistical parameters of the residuals were obtained after applying the multilinear regression of each analyzed dataset.

	1X	2X	5X	10X	30X	60X
residuals						
Mean[dBA]	0.0	0.0	0.0	0.0	0.0	0.0
Std [dBA]	1.021	1.011	1.010	0.992	0.963	0.967
skewness	0.559	0.617	0.559	0.559	0.634	0.629
kurtosis	0.504	0.881	0.807	1.015	1.389	1.456
Shapiro	0.979	0.977	0.981	0.979	0.975	0.976

It can be observed that, by enlarging the entries of the original dataset that are provided to the multilinear regressor, a corresponding enlarged number of residuals is generated that will populate the distribution. Since the distribution becomes sharper and sharper (as indicated by the kurtosis indexes, Table 2), it can be concluded that most of the residuals are in the center part of the distribution, making them less relevant than the ones located in the tail regions.

Such a shape also indicates that the multilinear regression technique applied to the dataset is valid since it generates residuals normally distributed and perfectly centered. The Skewness index of the residual populations presents no large fluctuation between the datasets, meaning that the symmetry of the populations is preserved whether the dataset is increased in size or not. Finally, the Shapiro-Wilk test for the assessment of the normality of the distribution indicates that all the populations of the residuals are normally shaped, having values bigger than 0.974. Figure 3 shows the distribution of the residual populations for all the datasets investigated. The investigation of the datasets proceeded with the

evaluation of the process of residual outliers removing, to test if it affected the calibration process.

As stated in section three, such a process consists of the removal from the original dataset of the entries that generate outliers of the distribution of the residuals, followed by a second repetition of the multilinear regression on the remaining data. Being the outliers of a distribution the values that are more distant from the mean value, such a process removes the entries of the dataset that more deviate from the average, and that negatively influence the multilinear regression. In such a way, we investigated whether the application of the multilinear regression to the remaining dataset results in a better calibration or not. To do so, a first analysis of which k_0 factor of equation 1 best fits the scope of improving the residuals population distribution of multilinear regression has been performed on the 1X dataset.

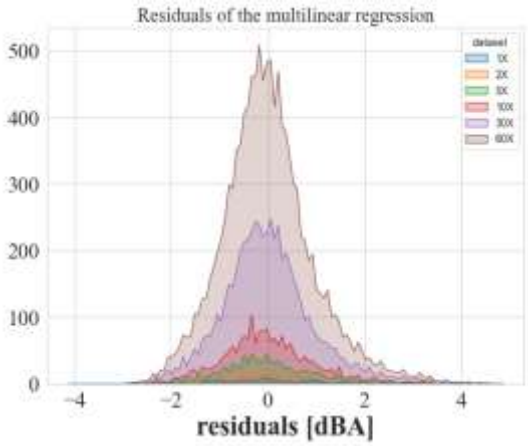

Fig. 3: Residuals of the multilinear regression technique on the generated datasets. The bigger the amount of data, the sharper the residual distribution, as also indicated by values reported in Table 2.

In such a way, we investigated whether the application of the multilinear regression to the remaining dataset results in a better calibration or not. To do so, a first analysis of which k_0 factor of equation 1 best fits the scope of improving the residuals population distribution of multilinear regression has been performed on the 1X dataset. The 1X dataset has been repeated 40 times by applying each time a different k_0 factor of equation 1, from 0.0 to 4.0, and the statistics of the obtained residuals have been analyzed. Such statistical residuals are reported in Table 3. The mean of the residual distribution is constantly 0.0, while the standard deviation of the residual significantly changes at varying of k_0 factor. In detail, when applying no outlier removal, the value of the standard deviation of the residuals is 1.021, which decreases at a minimum value of 0.417 when k_0 factor is equal to 0.1, in order to start increasing at the increase of k_0 value, up to the same value when no outliers removal procedure is applied.

The removal of residuals outliers, then, and the consequently second multilinear regression does actually improve the residuals population distribution. The best results are obtained, according to equation 1, by removing the data exceeding the mean for the ten percent of the interquartile range, while a wider range of exclusion is detrimental for the multilinear regression results. Skewness of the distribution of the residuals has a maximum value of 0.599 when no outliers removal is applied, then decreases to 0.323 at a value of k_0 equal to 0.1. As for the k_0 grows, the skewness value becomes bigger up to the same value of 0.599 when k_0 factor is equal to or bigger than 2.4. The symmetry of the distribution is then positively affected by the process of outliers removal. The kurtosis index is 0.504 without removing outliers, then decreases to -0.514 at k_0 equal to 0.1, and increases up to the original value.

Table 3. Statistical parameters of the 1X dataset residual population with and without the outliers removal process, at varying k_0 factor values

k_o	Mean [dBA]	Std [dBA]	Skewness	Kurtosis	Shapiro
None	0.0	1.021	0.559	0.504	0.979
0.1	0.0	0.417	0.323	-0.514	0.977
0.2	0.0	0.518	0.382	-0.658	0.969
0.3	0.0	0.598	0.205	-0.775	0.979
0.4	0.0	0.658	0.259	-0.737	0.975
0.8	0.0	0.783	0.367	-0.346	0.982
1.2	0.0	0.907	0.489	-0.025	0.976
1.6	0.0	0.990	0.507	0.246	0.980
2.0	0.0	0.990	0.507	0.246	0.980
2.4	0.0	1.021	0.599	0.504	0.979
2.8	0.0	1.021	0.599	0.504	0.979
3.2	0.0	1.021	0.599	0.504	0.979
3.6	0.0	1.021	0.599	0.504	0.979

Such behavior suggests that the sharpness of the distribution of the residuals is negatively affected by the removal of the outliers. However, it has to be considered that, even if the kurtosis index decreases, the corresponding standard deviation value is significantly lower compared to the original residual population (containing the outliers), making the process valid and worthy to be application. In Figure 4 the mean value of the residual is reported at varying of k_0 factor value (thick black line), together with a gray shaded area representing the standard deviation of the residuals themselves. As mentioned above, it is minimal when k_0 is equal to 0.1. The

same approach has been used to verify the combined effect of outliers removal and dataset increasing to test if the residual population could be further improved. To do so, the outliers removal approach has been applied to the 1X, 2X, 5X, 10X, 30X, and 60X datasets at k_0 value varying from 0.1 to 4.0.

Table 4 reports the results of the statistical analysis (in order to not overload the table clarity only results from the analysis of residuals at k_0 equal to 0.1, 0.4, 1.6, and 3.6 are reported). The mean value is constantly at 0.0, meaning that no matter which size of the dataset and whether the outliers removal procedure is applied or not, the distribution of the residuals is centered at 0.0. The standard deviation of the residuals always presents the same trend: it decreases at the increasing of the dataset size with the iterations and becomes bigger at varying the k_0 value, presenting the minimum value at k_0 equal to 0.1. The smallest value of the standard deviation of the residuals, then, is found when the outliers are removed with k_0 factor is equal to 0.1 is applied on the 60X dataset. Skewness presents a similar trend: it decreases with the increase of the dataset size but is limited to k_0 values between 0.1 and 0.4, where for higher k_0 values it suddenly increases. The best symmetry for the residuals is then obtained when the dataset is big and the outliers removal process is applied. The Kurtosis index does not change so sharply as it does when the whole dataset is considered (when no outliers are removed), while the Shapiro-Wilk test always reveals a normal distribution. Finally, from Table 4 it can be noted how, when the outliers removal process is applied, the main statistical parameters are identical for residuals of datasets 1X and 2X. The results of Table 4, relative to the standard deviation parameter, are also shown in Figure 5, where a heatmap permits the easy identification of the k_0 value and dataset size to minimize the standard deviation of the residual population. Dataset size and outliers removal process are the first two parameters studied in relation to the residual outcomes. A third aspect has been also investigated: the seed number. As stated before, the seed number is the specific value assigned to the in-built functions of Python to ensure the possibility of controlling the repetition of the experimental procedure: each time a seed value is assigned the function will return the same extracted values. To test the solidity of the written code and, in the end, of the built model, a test of the residual statistics over a dataset built with different seeds has been performed.

Table 4. Statistical parameters of the residuals population of all the analyzed datasets with and without the outliers removal process, at varying k_0 factor values

	1X	2X	5X	10X	30X	60X
No outliers removal						
Mean[dBA]	0.0	0.0	0.0	0.0	0.0	0.0
Std [dBA]	1.021	1.011	1.010	0.992	0.963	0.967
skewness	0.599	0.617	0.599	0.599	0.634	0.629
kurtosis	0.504	0.881	0.807	1.015	1.389	1.456
Shapiro	0.979	0.977	0.981	0.979	0.975	0.976
Outliers removal with k_0 factor = 0.1						
Mean[dBA]	0.0	0.0	0.0	0.0	0.0	0.0
Std [dBA]	0.417	0.417	0.378	0.357	0.347	0.348
skewness	0.323	0.323	0.188	0.162	0.104	0.087
kurtosis	-0.514	-0.514	-0.848	-0.932	-0.943	-0.905
Shapiro	0.977	0.977	0.978	0.975	0.978	0.981
Outliers removal with k_0 factor = 0.4						
Mean[dBA]	0.0	0.0	0.0	0.0	0.0	0.0
Std [dBA]	0.658	0.658	0.542	0.497	0.475	0.484
skewness	0.259	0.259	0.269	0.203	0.151	0.173
kurtosis	-0.737	-0.737	-0.642	-0.680	-0.713	-0.673
Shapiro	0.975	0.975	0.983	0.985	0.987	0.987
Outliers removal with k_0 factor = 1.6						
Mean[dBA]	0.0	0.0	0.0	0.0	0.0	0.0
Std [dBA]	0.990	0.990	0.910	0.868	0.826	0.830
skewness	0.507	0.507	0.370	0.334	0.326	0.314
kurtosis	0.246	0.246	0.157	0.193	0.265	0.199
Shapiro	0.980	0.980	0.988	0.989	0.990	0.990
Outliers removal with k_0 factor = 3.6						
Mean[dBA]	0.0	0.0	0.0	0.0	0.0	0.0
Std [dBA]	1.021	1.021	1.01	0.992	0.960	0.962
skewness	0.559	0.559	0.559	0.559	0.617	0.601
kurtosis	0.504	0.504	0.807	1.015	1.306	1.313
Shapiro	0.979	0.797	0.981	0.979	0.976	0.977

The same statistical parameters analyzed for a single dataset, then, have been compared between 100 different datasets of the same dimensions. Mean, standard deviation, kurtosis index, skewness, and value of the Shapiro-Wilk test have been collected and averaged over the 100 repetitions, to verify if the averaged values correspond to the mean values reported in Table 2. To prevent confusion and to facilitate the discussion, residuals from Table 2 come from a multilinear regression applied on datasets having seed=0, while datasets shown from now on, used for the assessment of the reliability of the model, have seed values from 1 to 100. The approach used for the following discussion, then, is to evaluate the averaged statistic values coming from datasets having seed 1 to 100 compared to the unique statistic values obtained from datasets having seed=0. From the results reported in Table 5, it is immediately visible how the change of seed does not affect the mean value of the residuals distribution, regardless of the dataset dimension and k_0 the value used to remove the outliers. For dataset 1X the averaged value, in fact, is 0.0 ± 0.0 and it is perfectly consistent with the dataset having seed=0.

Table 5. Statistical parameters of the residuals population of the analyzed datasets vs the averaged values of the residuals of the datasets repeated 100 times with different *seed*

	1X	2X	5X	10X	30X	60X
Residuals of datasets with *seed* = 0						
Mean[dBA]	0.0	0.0	0.0	0.0	0.0	0.0
Std [dBA]	1.021	1.011	1.010	0.992	0.963	0.967
skewness	0.599	0.617	0.599	0.599	0.634	0.629
kurtosis	0.504	0.881	0.807	1.015	1.389	1.456
Shapiro	0.979	0.977	0.981	0.979	0.975	0.976
Residuals of datasets with *seed* = 1 to 100 (averaged values)						
Mean[dBA]	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0
Std [dBA]	0.962 ± 0.069	0.973 ± 0.047	0.976 ± 0.027	0.977 ± 0.017	0.979 ± 0.012	0.979 ± 0.009
skewness	1.306 ± 0.869	1.364 ± 0.583	1.467 ± 0.457	1.501 ± 0.309	1.534 ± 0.167	1.554 ± 0.106
kurtosis	0.612 ± 0.233	0.607 ± 0.167	0.631 ± 0.107	0.631 ± 0.071	0.637 ± 0.044	0.637 ± 0.027
Shapiro	0.969 ± 0.015	0.973 ± 0.010	0.974 ± 0.007	0.974 ± 0.005	0.975 ± 0.003	0.975 ± 0.001

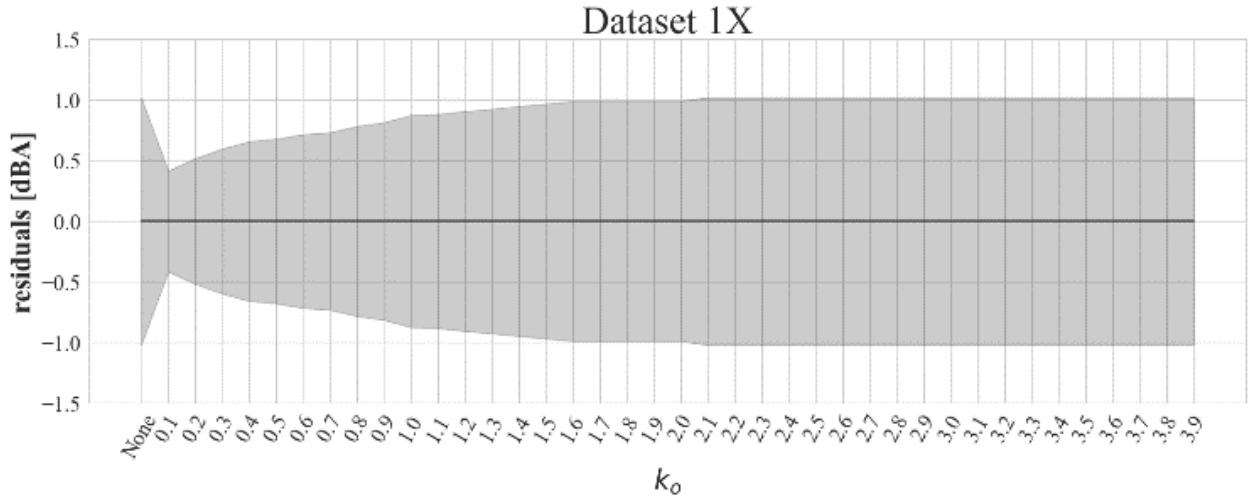

Fig. 4: Statistics of the 1X dataset residuals population with and without the outliers removal process, at varying k_0 factor values

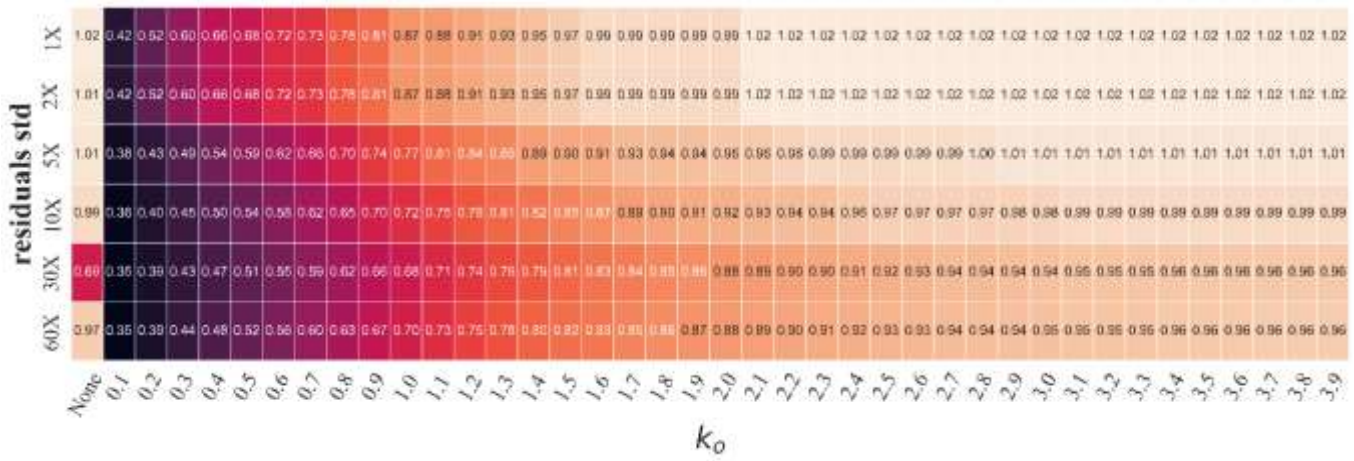

Fig. 5: heatmap correlating the dataset size and the k_0 factor applied to the process of outliers removal.

The standard deviation value is 0.962 ± 0.069, compared to the 1.021 of the dataset with seed=0. Shapiro Wilk test result is 0.969 ± 0.015, compared with 0.979 of seed=0, showing that all the residuals obtained populations are normally distributed.

Indices of asymmetry and sharpness of the residual distributions are also comparable, but the difference between the 100 repetitions is wider.

The Kurtosis index value for the 100 repetitions is 0.612 ± 0.233, where the one for the single dataset with seed = 0 is 0.599; the average skewness value of the 100 datasets is 0.612 ± 0.015 vs the one of a dataset at seed 0 equal to 0.504. Skewness and Kurtosis index are, in general, more variable over the 100 repetitions. At last, it is observable that a common trait of the repetition process of the datasets is that when the dataset is bigger in size, the wideness of the distribution of the averaged value is smaller. Overall, out of these results, it can be easily concluded that the random generation of the independent variables used to build the dataset is completely reliable, and it is not affected by the different repetitions of the process.

3.3 Validation of the Model

The investigations of the results regarding the validation of the model proceeded following the same criteria applied for the calibration: to verify if the different size of the generated dataset can improve the final results of the model (indicated by the error metrics) and if the final result changes at varying of the chosen seed. Since from the analysis of the calibration the most convenient value of k is shown to be 0.1, the following result analysis will focus on the comparison of the regression technique with and without outliers' removal only with the factor k=0.1. At first, the investigation of the multiregression coefficients is shown, since they are the ones used to simulate the final noise values. Table 6 reports the mean values of the coefficients at varying seeds, with and without outliers' removal process. The tables indicate that the values of the coefficients do not significantly change when applying the outlier's removal procedure. What is noticeable, nonetheless, is that the associated standard deviations reduced at increasing the dataset dimension (for all the coefficients). When calibrating the model using a larger dataset, then, the multiregression stage is performed with coefficient values more similar between the seeds, compared to when the multiregression is performed after a calibration with a shorter dataset. Consequently, the coefficients of the multi-regression will be more similar to each other when using a larger dataset, whichever seed has been chosen. Figure 6 reports the oscillation of the various coefficient values through the 100 chosen seeds when calibrating with a 1X and a 60X dataset, with and without outliers' removal.

Table 6. Mean values of the coefficients of the multilinear regression repeated 100 times with different *seeds*, with and without outliers' removal at *k=0.1*

	1X	2X	5X	10X	30X	60X
Multi-regression coefficients without outliers' removal						
Intercept	29.116 ± 1.667	29.071 ± 0.661	29.214 ± 1.146	29.117 ± 0.458	29.117 ± 0.293	29.180 ± 0.223
Coeff. Q	10.034 ± 0.167	10.024 ± 0.091	10.028 ± 0.128	10.023 ± 0.056	10.023 ± 0.033	10.022 ± 0.022
Coeff. vL	18.586 ± 0.646	18.588 ± 0.278	18.536 ± 0.411	18.612 ± 0.174	18.609 ± 0.123	18.607 ± 0.086
Coeff. vM	2.339 ± 0.520	2.360 ± 0.226	2.356 ± 0.390	2.346 ± 0.135	2.338 ± 0.093	2.336 ± 0.066
Coeff. vH	1.455 ± 0.513	1.477 ± 0.201	1.477 ± 0.327	1.451 ± 0.154	1.437 ± 0.103	1.440 ± 0.062
Coeff. P	2.363 ± 0.545	2.344 ± 0.254	2.326 ± 0.386	2.320 ± 0.177	2.311 ± 0.106	2.310 ± 0.071
Coeff. d	-25.646 ± 0.388	-25.643 ± 0.164	-25.658 ± 0.248	-25.638 ± 0.119	-25.641 ± 0.073	-25.641 ± 0.056
Multi regression coefficients with outliers' removal (*k = 0.1*)						
Intercept	28.383 ± 1.536	28.454 ± 1.107	28.332 ± 0.591	28.341 ± 0.402	28.376 ± 0.278	28.385 ± 0.210
Coeff. Q	10.031 ± 0.173	10.019 ± 0.129	10.016 ± 0.083	10.016 ± 0.048	10.015 ± 0.031	10.014 ± 0.022
Coeff. vL	19.042 ± 0.622	18.898 ± 0.433	19.023 ± 0.288	19.032 ± 0.170	19.030 ± 0.112	19.020 ± 0.080
Coeff. vM	2.190 ± 0.531	2.234 ± 0.395	2.222 ± 0.238	2.228 ± 0.131	2.221 ± 0.090	2.226 ± 0.067
Coeff. vH	1.449 ± 0.517	1.446 ± 0.339	1.469 ± 0.224	1.448 ± 0.160	1.450 ± 0.104	1.465 ± 0.064
Coeff. P	2.182 ± 0.552	2.151 ± 0.399	2.185 ± 0.252	2.168 ± 0.170	2.163 ± 0.109	2.159 ± 0.073
Coeff. d	-25.492 ± 0.391	-25.479 ± 0.245	-25.474 ± 0.162	-25.463 ± 0.115	-25.470 ± 0.072	-25.470 ± 0.050

To test if and how the validation process is also affected by the dataset size we finally performed the validation step with a dataset of all dimensions (1X,2X,5X,10X,30X,60X) and iterated the process through 100 seeds. The results are summarized in Figure 7. Dots represent the mean value of the errors of the validation process (identified as Measured $L_{eq,h}$ minus Simulated $L_{eq,h}$) through the 100 chosen seeds. The spread of the values constantly decreases at increasing the dimension of the original dataset, remaining on a mean value of -0.5 dBA. This means that the sensitivity of the model does not improve with increasing the size of the dataset, but it becomes more and more consistent over the seeds. In other words, when calibrating on a 30X or 60X original dataset, whichever random combination of values will produce a validation with an error value close to the mean one (in this application -0.5 dBA). By choosing a larger dataset for the calibration, then, the reproducibility is more likely than when using a smaller one. Figure 8 and Table 7 strengthen this aspect, by visualizing the mean error (averaged through the chosen 100 seeds) at different original dataset sizes (1X, 2X, 5X, 10X, 30X, 60X), with the

corresponding standard deviation. Whether the mean values do not move from -0.5 dBA, the standard deviation through seeds constantly reduces. It is important to underline that this aspect is not related to the final sensitivity of the model, i.e. the capability to simulate noise level as close as possible to the real measured ones, and that the two aspects are not related.

Table 7. Mean values (averaged through seeds) of the error of the validation step, by using different dataset sizes (without outliers' removal process)

1X	2X	5X	10X	30X	60X
-0.433 ± 0.249	-0.448 ± 0.156	-0.499 ± 0.100	-0.449 ± 0.081	-0.448 ± 0.048	-0.433 ± 0.033

To finally test the sensitivity of the model we compared the distribution of the values of Measured and Simulated $L_{eq,h}$ with different statistical parameters. We tested validation results for $L_{eq,h}$ values coming from calibration with and without outlier's removal, at 1X and 60X original datasets, as reported in Table 8. Mean values of Measured and Simulated are comparable, (72.085 vs. 72.491 for the 1X dataset; 72.085 vs 72.505 for the 60X dataset) as well as Skewness (-1.685 vs -1.195 for the 1X dataset; -1.685 vs -1.192 fr 60X dataset). The Kurtosis index differs between Measured and Simulated (4.872 vs 2.025 for the 1X dataset; 4.872 vs 2.023 for the 60X dataset), maybe because the model cannot correctly simulate values of the left tail of the Measured distribution (see Figure 9, where the two distributions are visually compared for dataset1X without outliers' removal). These values are probably due to anomalous situations of traffic, where the LTMS recorded unusually low $L_{eq,h}$ values. Shapiro-Wilk test results indicate that all the distributions are reasonably normal shaped (all the values are bigger than 0.885).

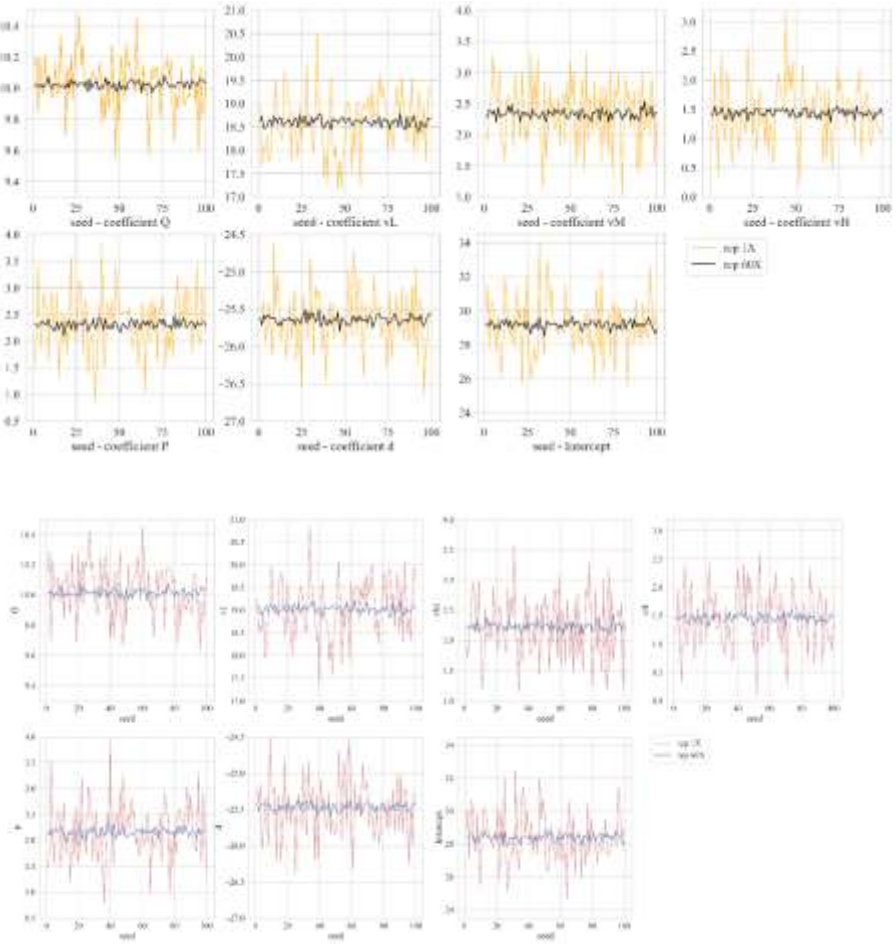

Fig. 6: Oscillation of the multi-regression coefficients over seed with (upper part, lines yellow and black) and without (lower part, lines red and blue) outliers' removal. In all the graphs the value of the coefficients derived from the application of the multiregression on a 1X dataset is compared to the values derived from the application of the multiregression on a 60X dataset.

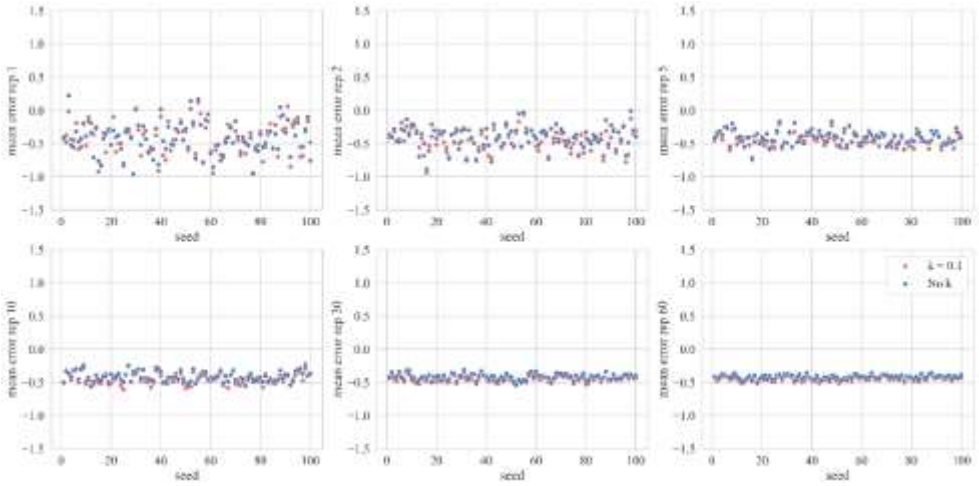

Fig. 7: Mean values of errors through the chosen seeds when calibrating with datasets of different sizes (1X,2X,5X,10X,30X,60X) with and without outliers' removal.

The last rows of Table 8 report error metrics MAE and MAPE. Also in this case there are no significant variations between the 1X dataset (MAE 1.747, MAPE 0.024 without outliers' removal, MAE 1.758, MAPE 0.025 with outliers' removal) and the 60X dataset (MAE 1.761, MAPE 0.023 without outliers' removal, MAE 1.759 MAPE 0.025 with outliers' removal). The findings summarized in Table 7, then, strengthen the observation that the dataset size and the process of outliers' removal do not influence positively nor negatively the final validation process in terms of error metrics.

Table 8. Statistical parameters of measured and simulated $L_{eq,h}$ values at 1X and 60X dataset, with and without outliers' removal (*k=0.1)*

	1X		**1X, *k= 0.1***		**60X**		**60X, *k= 0.1***	
	Meas	**Sim**	**Meas**	**Sim**	**Meas**	**Sim**	**Meas**	**Sim**
Mean	72.085 ± 1.999	72.491 ± 2.483	72.085 ± 1.999	72.508 ± 2.498	72.085 ± 1.999	72.541 ± 2.486	72.085 ± 1.999	72.505 ± 2.505
Skew	-1.685	-1.195	-1.689	-1.189	-1.685	-1.192	-1.685	-1.180
Kurt	4.872	2.025	4.872	1.987	4.872	2.023	4.872	1.995
Shapiro	0.886	0.923	0.886	0.924	0.886	0.924	0.886	0.924
MAE	1.747		1.758		1.761		1.759	
MAPE	0.024		0.025		0.023		0.025	

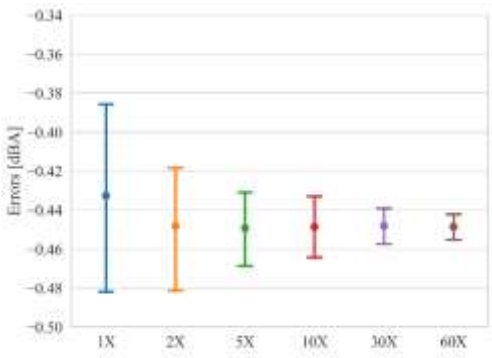

Fig. 8: Mean and standard deviations of error through 100 seeds when calibrating with datasets of different sizes (1X,2X,5X,10X,30X,60X)

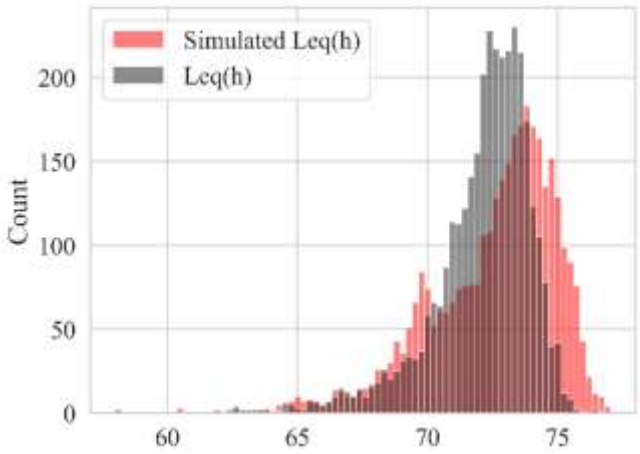

Fig. 9: Comparison of simulated and measured noise levels distributions are for dataset1X without outliers' removal

A last consideration has been made regarding the processing time used by the computer to return a final result of the calibration process and the validation process. As first two different times have been distinguished: the wall time and the CPU time. CPU time is the laps of time needed by the CPU to compute all the needed operations to return the results, whereas the wall time is the time effectively elapsed from the start of the calculation until the visualization of the results. CPU time can be affected by the contemporary execution of any other process in the background.

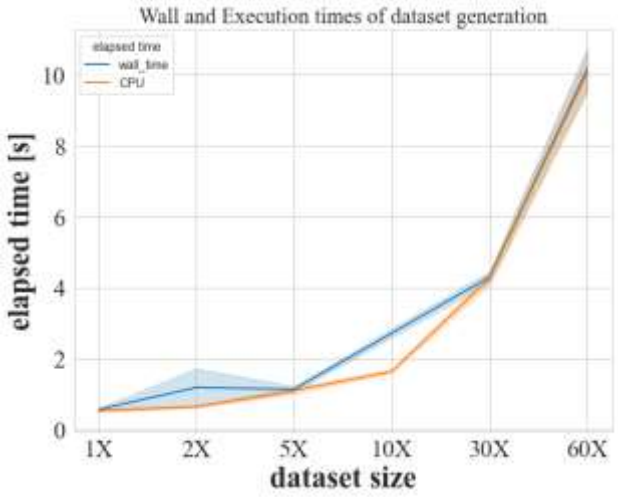

Fig. 10: CPU and Wall time of the calibration and validation of the model by using different sizes of the original dataset, expressed in seconds.

In the presented application no other relevant processes were running on the background, except for the essential ones. Wall time is generally higher than CPU time since it also involves the time needed by the compiler to process and visualize the results. Since the analyzed datasets are very different from each other in size, the time needed for their generation was studied. By comparing the execution times with the obtained results, in fact, a finer optimization of the model itself can be obtained. In Figure 10 a line is visualized, showing the total CPU and Wall time needed – expressed in seconds – for the calibration of datasets, depending on their dimension. CPU and Wall time needed for each dataset have been evaluated five times: the average value is shown, and the shaded area refers to the standard deviation. Please note that the elapsed time indicated in Figure 10 refers to the calibration and validation of datasets where no outlier removal has been performed. Figure 10 indicates that the time for the calibration of the model remains below 2 seconds until the dataset is 10X bigger than the standard one of 200 entries, but it doubles for the 30X dataset, and it doubles again for the 60X datasets. The elapsed time, then, significantly grows when the dataset gets bigger. By comparing Figure 5 with the results of Figure 10, anyway, it can be seen how the standard deviation of the residuals population does not present significant variation between 30X and 60X datasets (0.347 vs 0.348), making it not convenient for general prediction purposes when choosing a 60X datasets – requiring 10 s and a high percentage of CPU usage – to get a proper calibration process. Similarly, when comparing the outputs of the validation process (Figure 8 and Table 7), the difference, in terms of error, between 30X and 60X is not significant.

4 Conclusions

In this contribution, the reproducibility of the calibration and validation steps of a road traffic noise multilinear regression model based on a generated road traffic dataset is studied. The model has been previously described but here expanded and investigated in detail. The goodness of the calibration process has been estimated by analyzing in detail the statistical parameters of the residual population of the multilinear regression. The original dataset by which the model has been calibrated in previous works has been multiplied by specific factors to verify if an increased size of the input dataset could enhance the multilinear regression technique output. The size of the dataset has been observed to not greatly affect the residual statistical parameters. A second analysis has been performed by studying the effect of the multilinear regression technique on the removal of the outliers value of the residual population and of the application of a second identical multilinear regression technique. In such a case, even if the mean of the residual distribution has not changed, the wideness of the distribution itself is much smaller, meaning that the multilinear regression has been significantly improved. A combination of the two processes – increasing the size of the dataset and outliers removal – has also been tested, finding further improvement in the whole technique. Another analysis has been set on the dependence of the model on the input data, to test if the random generation of values of the independent variables could affect the final result of calibration. It has been observed that, in 100 different random generations of data, the final results are not affected. The aforementioned approach has been used to test the goodness of the validation step of the model, which has been tested on a road traffic dataset coming from a study of the Université Gustave Eiffel. The analysis revealed that, whether the mean value of the error metric is not affected by the outliers' removal process, it is strongly influenced by the size of the dataset where the calibration process is performed. By iterating the validation through the 100 seeds, in fact, it has been noted that the value of the mean error is greatly more stable on 60X datasets than on 1X datasets. It is safe to say, then, that when the calibration is applied on a larger generated dataset, the mean value of the final errors will be more stable and not depend on the random generation of the original dataset. To conclude, an

investigation on computing time of the generation-regression process has been analyzed at different dataset sizes, finding the best compromise between the time involved in the process and the final results in terms of mean value and wideness of the residuals and error populations. The here-tested model, then, correctly predicts $L_{eq,h}$ values with an MAE of 1.75 and a MAPE of 0.02, by using a generated dataset, which is particularly useful when real data are not available for the calibration. also present some criticalities, since it performs the calibration and validation only by using a single NEM, where the final results could be influenced by the usage of different strategies for computation of $L_{eq,h}$ values which will be used by the regressor: for this reason future works will focus on the implementing of different NEMs to test if the model will be stable and reliable and, consequently, suitable for a larger part of the scientific community.

Acknowledgement:
The authors acknowledge Antonio Pascale for the fruitful discussions about road traffic noise models.

References:
[1] European Environment Agency, *Environmental noise in Europe - 2020*, no. 22/2019. 2020.
[2] C. Clark, R. Crombie, J. Head, I. Van Kamp, E. Van Kempen, and S. A. Stansfeld, "Does traffic-related air pollution explain associations of aircraft and road traffic noise exposure on children's health and cognition? A secondary analysis of the United Kingdom sample from the RANCH project," *Am. J. Epidemiol.*, vol. 176, no. 4, pp. 327–337, 2012, doi: 10.1093/aje/kws012.
[3] D. Singh, N. Kumari, and P. Sharma, "A Review of Adverse Effects of Road Traffic Noise on Human Health," *Fluct. Noise Lett.*, vol. 17, no. 1, 2018, doi: 10.1142/S021947751830001X.
[4] M. Skrzypek, M. Kowalska, E. M. Czech, E. Niewiadomska, and J. E. Zejda, "Impact of road traffic noise on sleep disturbances and attention disorders amongst school children living in upper Silesian Industrial Zone, Poland," *Int. J. Occup. Med. Environ. Health*, vol. 30, no. 3, pp. 511–520, 2017, doi: 10.13075/ijomeh.1896.00823.
[5] D. Ouis, "Annoyance from road traffic noise: A review," *J. Environ. Psychol.*, vol. 21, no. 1, pp. 101–120, 2001.
[6] D. Petri, G. Licitra, M. A. Vigotti, and L. Fredianelli, "Effects of exposure to road, railway, airport and recreational noise on blood pressure and hypertension," *Int. J. Environ. Res. Public Health*, vol. 18, no. 17, 2021, doi: 10.3390/ijerph18179145.
[7] L. Fredianelli, M. Bolognese, F. Fidecaro, and G. Licitra, "Classification of noise sources for port area noise mapping," *Environ. - MDPI*, vol. 8, no. 2, pp. 1–16, 2021, doi: 10.3390/environments8020012.
[8] A. Mascolo, D. Rossi, A. Pascale, S. Mancini, M. C. Coelho, and C. Guarnaccia, "Noise Assessment during Motor Race Events: New Approach and Innovative Indicators," *WSEAS Trans. Environ. Dev.*, vol. 19, pp. 80–88, 2023, doi: 10.37394/232015.2023.19.7.
[9] P. Fernandes *et al.*, "Impacts of roundabouts in suburban areas on congestion-specific vehicle speed profiles, pollutant, and noise emissions: An empirical analysis," *Sustain. Cities Soc.*, vol. 62, Nov. 2020, doi: 10.1016/j.scs.2020.102386.
[10] W. Gardziejczyk and M. Motylewicz, "Noise level in the vicinity of signalized roundabouts," *Transp. Res. Part D Transp. Environ.*, vol. 46, pp. 128–144, 2016, doi: 10.1016/j.trd.2016.03.016.
[11] R. A. Hood, "Accuracy of calculation of road traffic noise," *Appl. Acoust.*, vol. 21, no. 2, pp. 139–146, 1987.
[12] Der Bundesminister für Verkehr Abteilung Straßenbau, "Richtlinien für den Lärmschutz an Straßen RLS-90," 1990.
[13] G. Dutilleux, J. Defrance, B. Gauvreau, and F. Besnard, "The revision of the French method for road traffic noise prediction," *Proc. - Eur. Conf. Noise Control*, no. July, pp.875-880, 2008.
[14] S. Sakamoto, "Road traffic noise prediction model '"ASJ RTN-Model 2018"': Report of the research committee on road traffic noise," *Acoust. Sci. Technol.*, vol. 41, no. 3, pp.529-589, 2020,.
[15] F. A.-L. S. Kephalopoulos, M. Paviotti, *Common Noise Assessment Methods in Europe (CNOSSOS-EU)*, no. 2. 2012, [Online], http://europa.eu/ (Accessed Date: November 2, 2023).
[16] D. Rossi, A. Mascolo, and C. Guarnaccia, "Calibration and Validation of a Measurements-Independent Model for Road

Traffic Noise Assessment," *Appl. Sci.*, vol. 13, no. 10, 2023.

[17] B. Gauvreau, "Long-term experimental database for environmental acoustics," *Appl. Acoust.*, vol. 74, no. 7, pp. 958–967, 2013.

[18] R. L. Wayson, T. W. A. Ogle, and W. Lindeman, "Development of Reference Energy Mean Emission Levels for Highway Traffic Noise in Florida," *Transportation Research Record*, 1416. pp. 82–91, 1993.

[19] J. Quartieri, G. Iannone, and C. Guarnaccia, "On the improvement of statistical traffic noise prediction tools," *Proc. 11th WSEAS Int. Conf. Acoust. Music Theory Appl. AMTA '10*, no. June, pp. 201–207, 2010.

Contribution of Individual Authors to the Creation of a Scientific Article (Ghostwriting Policy)

Conceptualization: D. Rossi, C. Guarnaccia
Data curation: A. Mascolo, D. Rossi
Investigation: D. Rossi, C. Guarnaccia
Methodology: D. Rossi, C. Guarnaccia
Software: D. Rossi
Supervision: C. Guarnaccia
Visualization: A. Mascolo, D. Rossi
Writing - original draft: A. Mascolo, D. Rossi
Writing - review & editing: A. Mascolo, D. Rossi, C. Guarnaccia

Sources of Funding for Research Presented in a Scientific Article or Scientific Article Itself

No funding was received for conducting this study.

Conflict of Interest

The authors have no conflicts of interest to declare.