# Development and Application of a Time Series Predictive Model to Acoustical Noise Levels

Claudio Guarnaccia*, Joseph Quartieri*, Nikos E. Mastorakis+, Carmine Tepedino*

*Department of Industrial Engineering, University of Salerno
Via Giovanni Paolo II, I-84084 Fisciano (SA) – ITALY
cguarnaccia@unisa.it , quartieri@unisa.it , ctepedino@unisa.it

+Technical University of Sofia, English Language Faculty of Engineering
Industrial Engineering, Sofia 1000, Sofia – BULGARIA
http://www.wseas.org/mastorakis

*Abstract:* - Acoustical noise is one of the most relevant problem to be faced in urban areas. Since a large network of monitoring stations would be very expensive, the adoption of predictive model is a very common practice in environmental impact assessment. Usually, the standard models are based on the fit of field measurements in certain areas, in order to evaluate the model parameters and be able to give predictions in any other condition. A new approach, based on Time Series analysis, is presented in this paper. This approach considers that a time series can be composed by a trend (long term behaviour), a seasonality (periodicity) and an irregular term (random non deterministic variation). The evaluation of these components is made in the calibration phase and a validation can be performed evaluating the difference between observed and forecasted values. This scheme is applied to a case study time series, that is a daily noise levels dataset detected in the city of Messina, Italy. The building model procedure and results are discussed and presented in details for different calibration and validation subsets, in order to highlight the variation in the model predictive capabilities and to optimize the forecast, by means of minimization of the difference between actual data and forecasts. The easiness in the implementation and the good predictive performances will be the strengths of the model.

*Key-Words:* - Time Series, Acoustics, Noise Control, Predictive Model.

## 1 Introduction

Urban agglomerations are quickly growing, according to industrial, logistics and transportation optimization [1]. This growth leads to environmental problems that need to be carefully threated, such as air pollution, acoustic noise, electromagnetic field, etc. In [2], the authors proposed a method to include several pollution agents in an unique complex index. More in general, environmental impact must be taken into account when defining urban planning policy, because of the risk related to noise exposure (see for instance [3]).

Regarding acoustic noise, it must be considered that it has a highly random nature and it is difficult to be described, in terms of source, propagation and persistence. Thus, it is necessary on one hand to have a large network of measurement apparatus, or, on the other hand, to adopt advanced mathematical, statistical and probabilistic tools. The strong need for monitoring of acoustic emissions, in fact, is competing with the high costs of installing and maintenance of sound level meters and related equipment for long term acquisition. This is the reason why several predictive models have been developed in literature. For instance, regarding noise produced by vehicular traffic flow, the first models appeared in the ’60, and were typically based on regressive models. In [4], a detailed comparison between several models adopted by different countries regulations is presented, while in [5] the models predictions are compared with experimental data. The example of traffic noise models shows that it is very difficult to obtain a general statistical model (based on measured data), able to give predictions to be used in different countries and environments, with different boundary conditions. In [6-12] and references therein, these difficulties are largely discussed and dynamical approaches are suggested to overcome the shortcomings of standard models. In addition, the standard models very often adopt a hourly time base, that sometimes is not easily extendable to the daily noise levels.

In this paper, the authors present a very different approach, based on a mathematical method, the Time Series analysis (TSA) (see for instance [13, 14]), able to reproduce the behaviour of the data series and to give predictions for future values of acoustic noise. The analysis is performed on a large data set of acoustic noise measurements taken in a city of South Italy, Messina. In this case study, the noise is mainly due to road traffic flow and it evidences very interesting features, in terms of periodicity and low variability.

The reliability of the model will be discussed in terms of displacement between measured and predicted values. In addition, an analysis of the error of the model as a function of calibration data set size will furnish interesting indications for model optimization.

## 2 Methods

Time Series analysis (TSA) models are mathematical models able to reproduce the slope of a certain data series and to forecast future values. These models are largely used in several disciplines, such as Economics, Physics, Engineering, Mathematics, etc.. (see for instance [15-17]).

TSA have two general aims: first of all, the identification of the intrinsic features of the considered data, representing the phenomenon under study; then, the possibility to predict the future behaviour of the observed data series. There exist different approaches, many of them deeply studied in literature (see for example [18]).

TSA models are mostly adopted when the data sets follow recurring seasonal patterns. Thus, a general procedure may be resumed as follows:

- Eventual seasonal effect detection in the data set
- Lag (periodicity) evaluation
- Smoothing (removal of periodicity) of the data time series
- Trend and seasonality evaluation
- Final model drawing

The presence of a periodicity in the time series may be confirmed by means of autocorrelation coefficient evaluation:

$$\rho(k) = \frac{\sum_{t=1}^{n-k}(x_t-\bar{x}_{s1})(x_{t+k}-\bar{x}_{s2})}{\sqrt{\sum_{t=1}^{n-k}(x_t-\bar{x}_{s1})^2}\sqrt{\sum_{t=1}^{n-k}(x_{t+k}-\bar{x}_{s2})^2}} \qquad (1)$$

where $x_t$ is the measurement value in $t$, $n$ is the number of periods, $k$ is the lag and

$$\bar{x}_{s1} = \frac{\sum_1^{n-k} x_t}{n-k} \quad ; \quad \bar{x}_{s2} = \frac{\sum_1^{n-k} x_{t+k}}{n-k} \qquad (2)$$

The maximization of autocorrelation is a useful instrument to evaluate the lag value. Then, the trend of the data may be studied by means of regressive tools, for example linear regression, applied on actual measurements or on the moving average data set (data not influenced by seasonality).

The prediction may be provided by adding or multiplying the components of the forecast, defining the “additive” or “multiplicative” TSA model.

The general theoretical assumption of the TSA model presented in this paper is that the random variable $A_t$, at any period $t$, is given by:

$$A_t = T_t\,\bar{S}_i + e_t \qquad (3)$$

where $T_t$ the trend, $\bar{S}_i$ the seasonal effect coefficient (defined below) and $e_t$ is the irregular component, not deterministically predictable. Let us underline that the period index $t$ varies from 1 to $n$, the total number of periods, while $i$ varies from 1 to $k$, the lag coefficient, assuming that the seasonal effect is periodic. In particular, as will be explained below, for a given $t$, if $t<k$, $i$ is the remainder of the ratio between $t+k$ and $k$; if $t=k$, then $i=k$; if $t>k$, $i$ is the remainder of the ratio between $t$ and $k$. For instance, in this case study, assuming a weekly periodicity, $i$ will vary from 1 to 7, and will represent each day of the week (Monday, Tuesday, etc.).

A first formula for the punctual prediction of the model, $F_t$ , is given by the multiplication between the trend and the seasonality:

$$F_t = T_t\,\bar{S}_i \qquad (4)$$

Therefore, the model considers that the point forecast at a given period is the combination of a trend component, i.e. long term measurements behaviour, and a correction due to the specific period in which the prediction is made.

In this paper, the authors adopted a moving average technique to remove the seasonality of the data series. The choice was a centred moving average, with width given by the lag. Our data set suggested a lag of 7 periods (days), according to a weekly periodicity of the traffic noise emission. This lag was validated by autocorrelation coefficient estimation, as reported in the next section.

Then, the seasonal effect $S_t$ at a given period $t$, is obtained by the ratio between the actual data $A_t$ and the moving average value $M_t$ . Remember that for a dataset of $n$ periods, $n - k + 1$ values are available for the centred moving average and, consequently, for the seasonality effect $S_t$.

$$S_t = \frac{A_t}{M_t} \qquad (5)$$

A seasonal coefficient $\bar{S}_i$, evaluated on all the homologous periods, was estimated by averaging the seasonal effect, as follows:

$$\bar{S}_i = \frac{\sum_{l=0}^{m_i-1} S_{i+lk}}{m_i} \tag{6}$$

where $m_i$ is the number of homologous *i*-esim periods (in our case, the number of Mondays, Tuesdays, etc.) in the total time range.

The trend component has been calculated by means of a linear regression model:

$$T_t = b_0 + b_1 t \tag{7}$$

The coefficients have been evaluated on the centred moving average, defined above.

The term $e_t$ is considered to take into account irregularity of the data series and it may be evaluated in the calibration of the model, i.e. when the measurements are available, as the subtraction between actual data and forecasted values at a given period $t$ :

$$e_t = A_t - F_t \tag{8}$$

Assuming that the irregular term is normally distributed, its mean coincides with the mode, i.e. the most likely value. Thus, the mean can be added to the forecast, in order to improve the model prediction.

The final results is a mixed TSA model, i.e. multiplicative between trend and seasonality and additive for the irregularity.

Let us remind that this TSA model can be calibrated (estimation of model parameters) on a given time interval, according to the procedure described above, and then it can be validated on a successive range of data, not used in the calibration phase, comparing the forecasted values with the actual measurements in each period. In the next section, a comparison between different calibration and validation range choices will be presented.

In order to estimate the effectiveness of the model, the statistical features of the difference between actual and point forecasted value (error) have been studied, both in the calibration and in the validation phases. A frequencies histogram of the errors is presented, together with the statistics, that are mean, standard deviation, median, min and max. In addition, skewness and kurtosis indexes are calculated to evaluate the normality of the error distributions.

Quantitative metrics of error are given by the "Mean Percentage Error" (MPE) and "Coefficient of Variation of the Error" (CVE), according to the following formulas:

$$MPE = \frac{\sum_{t=1}^{n}\left(\frac{A_t - F_t}{A_t}\right)100}{n} \tag{9}$$

$$CVE = \frac{\sqrt{\frac{\sum_{t=1}^{n}(e_t)^2}{n-1}}}{\bar{A}} \tag{10}$$

where $\bar{A}$ is the mean value of the actual data in the considered time range.

Formula (9) (MPE) gives a measurement of the error distortion, while formula (10) (CVE) furnishes the error dispersion. MPE is able to describe if the model overestimates or underestimates the reality, while CVE considers the variation from the actual value in absolute value.

## 3 Data Analysis and Results

The data set used in this paper is related to a field measurement long term campaign designed and performed by the local government of Messina, a city in the South of Italy. Messina has about 240000 inhabitants and, besides the usual air pollution problems of a medium size city with a commercial dock, a very high traffic flow and several industrial settlements, it has a relevant noise pollution, mainly due to transportation infrastructures. The local administration decided to place several monitoring stations, equipped with first class sound level meters, in order to measure the long term fluctuation of noise levels. These data have been made available by the local administration on a web platform [19].

The authors focused on the monitoring station of Viale Boccetta, considering the A weighted equivalent level [20, 21], defined as follows:

$$L_{Aeq,T} = 10 \log\left[\frac{1}{T}\int_0^T \frac{p_A^2(t)}{p_0^2}dt\right] \tag{11}$$

related to a daily interval of 16 hours, from 6 a.m. to 10 p.m.. The time range goes from the 11th of May 2007 to the 26th of March 2008, i.e. 321 days/periods.

The summary statistics of complete data set are resumed in Tab. 1.

**Tab. 1:** Summary of statistics of the complete data set.

| Mean [dBA] | Std.dev [dBA] | Median [dBA] | Min [dBA] | Max [dBA] |
|---|---|---|---|---|
| 73.07 | 0.65 | 73.5 | 70.5 | 75.0 |

Since the data are strongly related to vehicular traffic flows, an evident seasonal effect is present.

A lag of 7 days has been supposed, according to a weekly seasonality. The choice seems reasonable looking at the data time series (Fig. 1) and considering that during the weekend a lower traffic flow is observed. The evaluation of autocorrelation coefficient gave a value of 0.58.

The trend has been obtained calculating the mobile average (removal of the seasonality) and evaluating the linear regression (see Fig. 2).

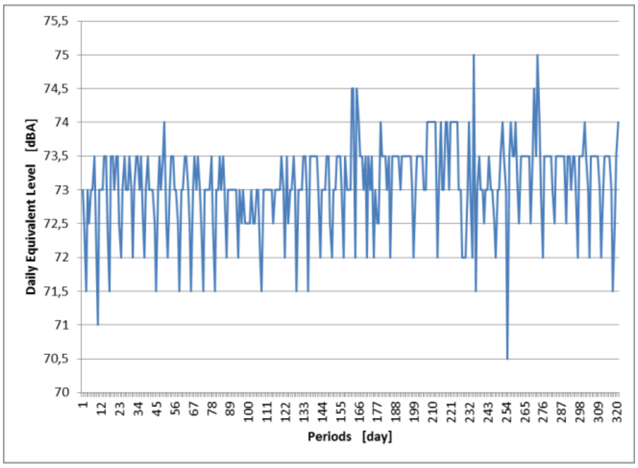


**Fig. 1:** Time series of the daily equivalent levels in the complete range, from the 11th of May 2007 to the 26th of March 2008, i.e. 321 days/periods.

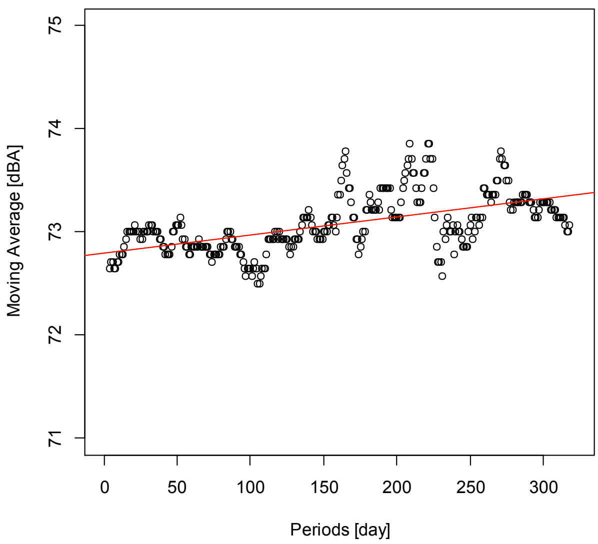


**Fig. 2:** Centred moving average, with width equal to 7, of the daily equivalent level in the complete time range. The red line is the linear regression.

### 3.1 Calibration on the first 150 data and validation on the following 171 data

The first analysis has been performed calibration the model parameters on the first 150 data and validating the resulting model on the remaining 171 data of the series.

During the calibration phase, the error (difference between actual and forecasted values) frequencies have been plotted in a histogram (Fig. 3). The evaluation of the skewness and kurtosis confirms the hypothesis of normal distribution for the errors.

The tuned model has been used to predict the data in the remaining part of the time series (171 data). In particular, a prediction interval has been fixed, assuming a half width of 2 standard deviations ($s_e$):

$$PI = F_t + m_e \pm 2s_e \qquad (12)$$

where $m_e$ is the mean error evaluated (such as $s_e$) on the calibration data set.

These results are resumed in Figure 4, where the blue line represents the actual data, the red one is the point forecast of the model (centre of the prediction interval) and the purple dashed lines are the lower and upper bounds of the interval.

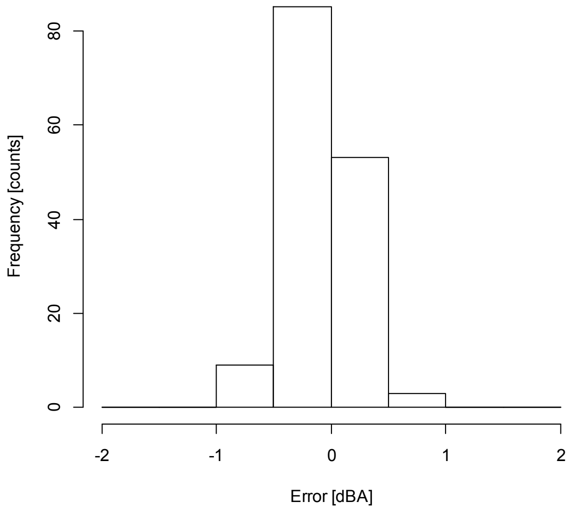


**Fig. 3:** Frequency histogram of the errors calculated on the model calibration, performed on the first 150 data.

**Tab. 2:** Summary of statistics of the error distribution evaluated on the calibration.

| Mean [dBA] | Std.dev [dBA] | Median [dBA] | Min [dBA] | Max [dBA] | skew | kurt |
|---|---|---|---|---|---|---|
| 0.00 | 0.30 | -0.09 | -0.81 | 0.72 | -0.03 | -0.10 |

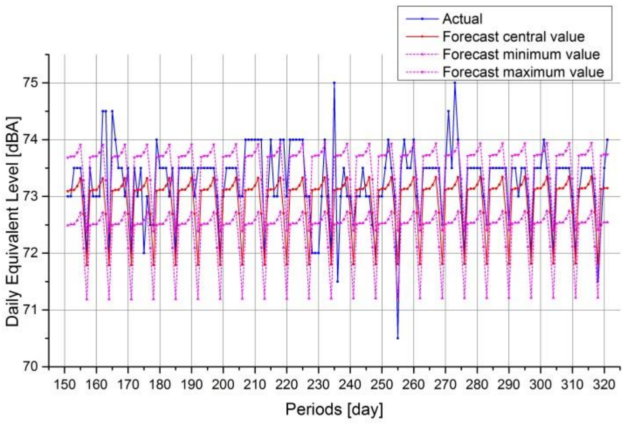


**Fig. 4:** Prediction interval of the model. Blue line represents the actual data, the red one is the average prediction of the model and the purple dashed lines are the lower and upper bounds of the interval.

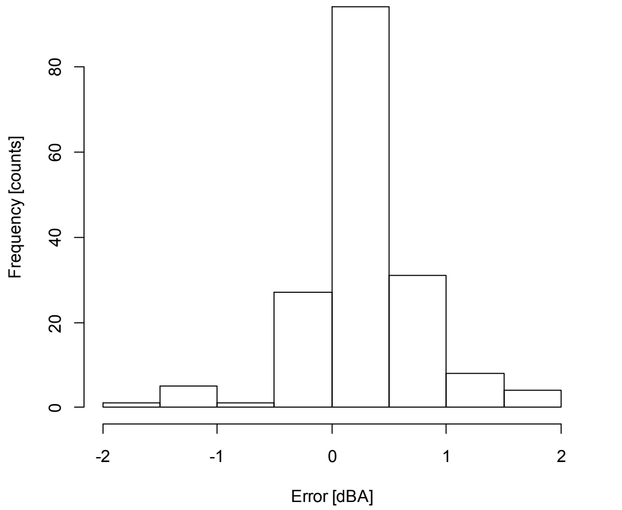


**Fig. 5:** Frequency histogram of the errors calculated on the model validation, performed on the latter 171 data.

**Tab. 3:** Summary of statistics of the error distribution evaluated on the validation.

| Mean [dBA] | Std.dev [dBA] | Median [dBA] | Min [dBA] | Max [dBA] | skew | kurt |
|---|---|---|---|---|---|---|
| 0.33 | 0.52 | 0.32 | -1.63 | 1.89 | -0.29 | 2.57 |

The validation on the following 171 data furnished good results, as shown by the statistics of the error distribution, resumed in Table 3. A mean error of 0.33 dBA evidences a little underestimation of the model, since the error is defined as the actual data minus the forecast in each period. The standard deviation of about 0.5 dBA gives a confidence interval half width of about 1 dBA.

### 3.2 Calibration on increasing data size and validation on 3 different 50 data intervals

The next analysis consists in a comparison between the results obtained increasing the calibration data size and validating the model on the following 50 data.

The calibration datasets and the respective validation ranges are reported in Fig. 6.

The plots of the model forecast in 3 different cases are reported in Figures 7, 8 and 9, while in Table 4 the main statistics obtained in the 3 validations are resumed.

Figures 7, 8 and 9 show that there is generally an agreement between model forecasts and actual data, especially after the third step (Fig. 9), with the largest calibration data size. This result is confirmed by statistics of each analysis, resumed in Table 4. The first step of calibration data size increase (from 150 to 200 data) produces a lowering of the mean error but a growth of the standard deviation. In the last analysis, i.e. a calibration data size of 250 periods, the mean is approximately the same of the previous step, while the standard deviation decreases as expected.

The error distribution histograms corresponding to the three validation intervals are reported in Figures 10, 11 and 12. The assumption of normality for error distribution is more or less respected changing the calibration dataset size and shifting the validation period, except for the second case. These results on mean error and standard deviation, resumed in Fig. 13, can be explained considering the statistics, in particular the spread of data, of each validation dataset (Table 5). The standard deviation, in fact, shows a slight growth in the second validation dataset (when the model is calibrated on 200 data), resulting in a more variable curve. It is important to notice that if the standard deviation of the data in the validation dataset increases, the assumption of normal distribution of error is easily violated.

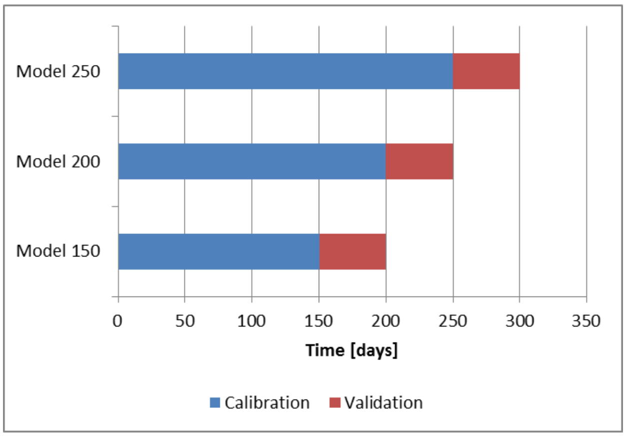


**Fig. 6:** Increasing calibration datasets and respective validation ranges.

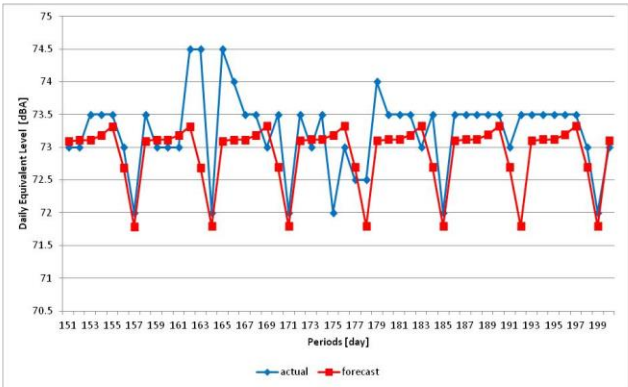


**Fig. 7:** Comparison between the forecasts, obtained tuning the model on the first 150 data, and the following 50 actual data.

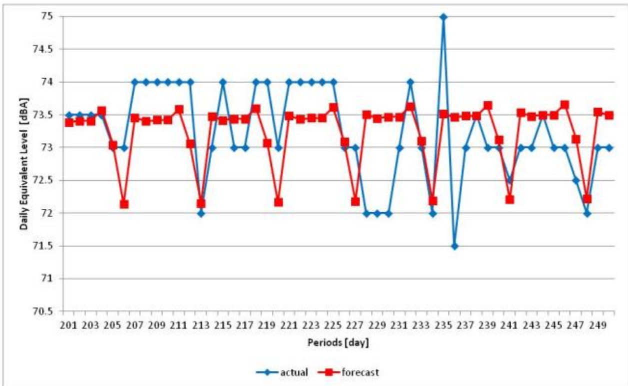


**Fig. 8:** Comparison between the forecasts, obtained tuning the model on the first 200 data, and the following 50 actual data.

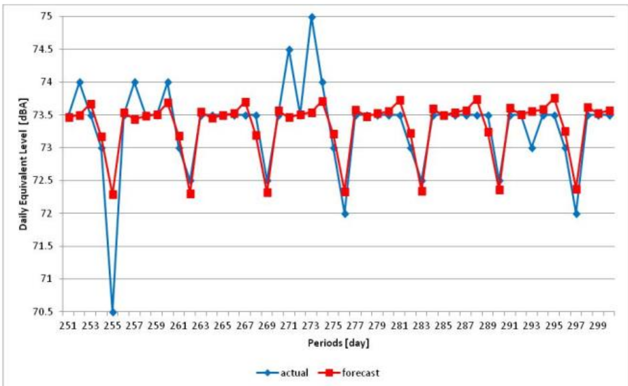


**Fig. 9:** Comparison between the forecasts, obtained tuning the model on the first 250 data, and the following 50 actual data.

**Tab. 4:** Comparison between statistics of the error distributions as a function of calibration data size, in the validation phase.

| Tuning data size | Mean [dBA] | Std.dev [dBA] | Median [dBA] | Min [dBA] | Max [dBA] |
|---|---|---|---|---|---|
| 150 | 0.32 | 0.51 | 0.31 | -1.19 | 1.81 |
| 200 | -0.02 | 0.70 | -0.02 | -1.96 | 1.48 |
| 250 | -0.01 | 0.42 | -0.04 | -1.79 | 1.46 |

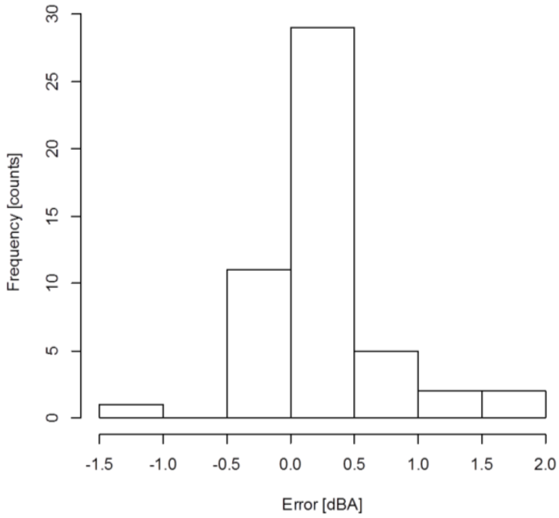


**Fig. 10:** Frequency histogram of the errors calculated on the model validation, performed on the data from 151st to the 200th.

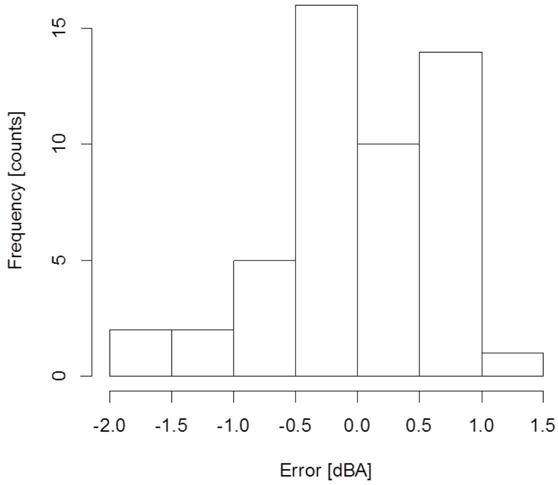


**Fig. 11:** Frequency histogram of the errors calculated on the model validation, performed on the data from 201st to the 250th.

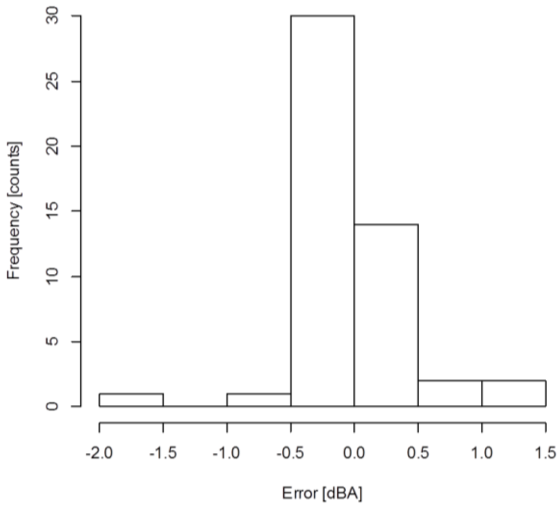


**Fig. 12:** Frequency histogram of the errors calculated on the model validation, performed on the data from 251st to the 300th.

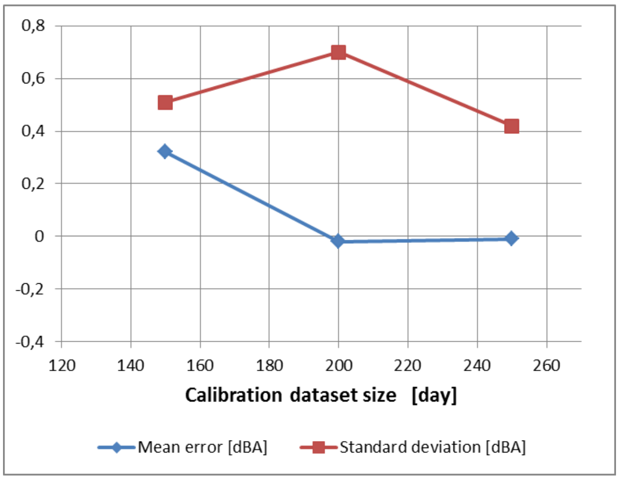


**Fig. 13:** Mean and standard deviation of the error distribution obtained validating the model on 50 actual data, as a function of the calibration data size.

**Tab. 5:** Summary statistics of validation data used to test models tuned with different data size.

| Validation data | Mean [dBA] | Std.dev [dBA] | Median [dBA] | Min [dBA] | Max [dBA] |
|---|---|---|---|---|---|
| 151-200 | 73.23 | 0.62 | 73.5 | 72.0 | 74.5 |
| 201-250 | 73.23 | 0.74 | 73.0 | 71.5 | 75.0 |
| 251-300 | 73.33 | 0.67 | 73.5 | 70.5 | 75.0 |

**Tab. 6:** MPE and CVE (error metrics) values, calculated in the calibration and validation phases, for different calibration data set size.

| | Calibration data size | | |
|---|---|---|---|
| | 150 | 200 | 250 |
| MPE (calibration) | -0.00697 | -0.01818 | -0.02172 |
| MPE (validation) | 0.43773 | -0.03565 | -0.02035 |
| CVE (calibration) | 0.004097 | 0.005047 | 0.006159 |
| CVE (validation) | 0.008246 | 0.009509 | 0.005691 |

Finally, the prediction errors MPE and CVE defined in section 2, respectively formula (9) and formula (10), are calculated in the three steps described above. Results are reported in Table 6.

### 3.3 Calibration dataset size increasing to the past and validation on the same 50 observed data

In order to check the model capability to predict future measurements, the authors performed the calibration, enlarging the dataset, on three different time ranges: starting from the 101$^{st}$ day to the 250$^{th}$ (150 days), from the 51$^{st}$ to the 250$^{th}$ (200 days) and from the 1$^{st}$ to the 250$^{th}$ period, considering all the data available before the validation period, i.e. from the 251$^{st}$ to the 300$^{th}$ day (50 days).

The calibration datasets and the fixed validation range are reported in Fig. 14. The model parameters estimated using the three different data ranges are reported in Table 7.

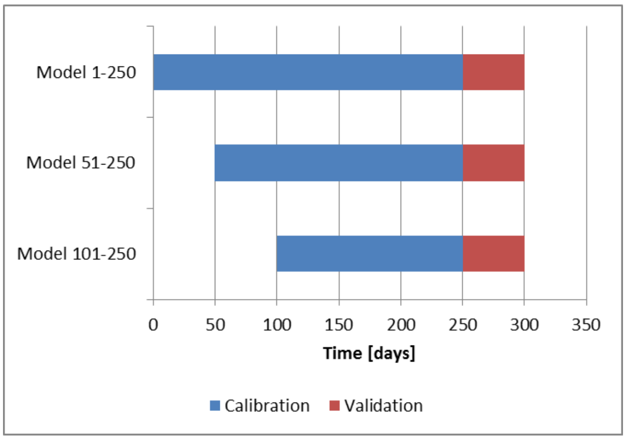


**Fig. 14:** Increasing calibration datasets and fixed validation range.

**Tab. 7:** Model parameters estimated using the three different data ranges.

| | Calibration ranges [days] | | |
|---|---|---|---|
| | 101-250 | 51-250 | 1-250 |
| $b_0$ | 0.002804236 | 0.00295778 | 0.00205197 |
| $b_1$ | 72.93012543 | 72.763531 | 72.7738095 |
| $\bar{S}_1$ monday | 1.00380 | 1.00354 | 1.00321 |
| $\bar{S}_2$ tuesday | 1.00038 | 1.00101 | 1.00184 |
| $\bar{S}_3$ wednesday | 1.00171 | 1.00248 | 1.00243 |
| $\bar{S}_4$ thursday | 1.00224 | 1.00248 | 1.00280 |
| $\bar{S}_5$ friday | 1.00484 | 1.00443 | 1.00514 |
| $\bar{S}_6$ saturday | 0.99957 | 0.99908 | 0.99826 |
| $\bar{S}_7$ sunday | 0.98741 | 0.98707 | 0.98625 |

Examining model coefficients is possible to obtain a lot of information about the behaviour of the pollutant agent under study. In our case the noise levels in the area subjected to measurement do not show a relevant variation in the trend: the $b_0$ slope of the regression line is close to zero. Thus, the $b_1$ coefficient can be read as the mean value of the noise level in the substantially short period of time under analysis. This type of pollutant agent seems to have a fast response to the human activities: the seasonal coefficients are lower than one in Saturday and Sunday, i.e. during the weekend, where there is a decreasing in the noise sources like traffic flow and activity in the commercial dock near the measurement station.

The results of the three models are reported in Fig. 15, 16 and 9. Let us remind that the last model application, calibrated on the dataset from the 1$^{st}$ to the 250$^{th}$ period and validated on the last 50 days (from the 251$^{st}$ to the 300$^{th}$), is the same as in subsection 3.2.

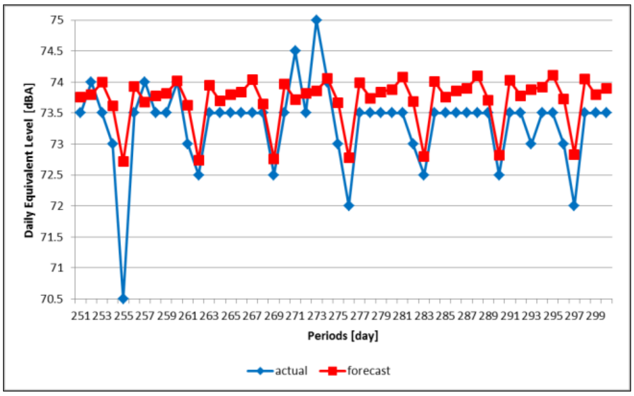


**Fig. 15:** Comparison between the forecasts, obtained calibrating the model on the data from the 101st to the 250th. The comparison is on the 50 actual data from the 251st to the 300th.

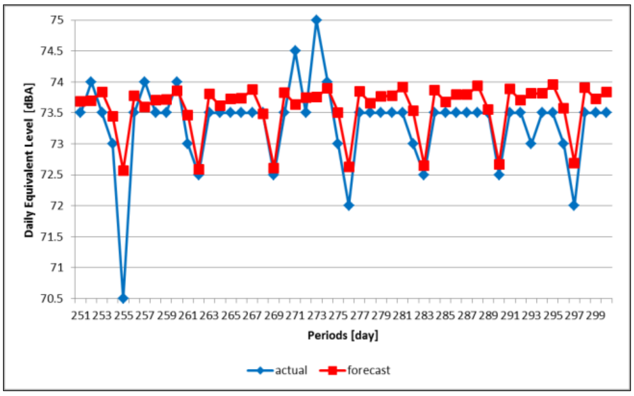


**Fig. 16:** Comparison between the forecasts, obtained calibrating the model on the data from the 51st to the 250th. The comparison is on the 50 actual data from the 251st to the 300th.

**Tab. 8:** Comparison between statistics of the error distributions in the validation phase as a function of calibration range.

| Calibration range | Mean [dBA] | Std.dev [dBA] | Median [dBA] | Min [dBA] | Max [dBA] |
|---|---|---|---|---|---|
| 101-250 | -0.37 | 0.44 | -0.37 | -2.22 | 1.14 |
| 51-250 | -0.25 | 0.43 | -0.28 | -2.07 | 1.25 |
| 1-250 | -0.01 | 0.42 | -0.04 | -1.79 | 1.46 |

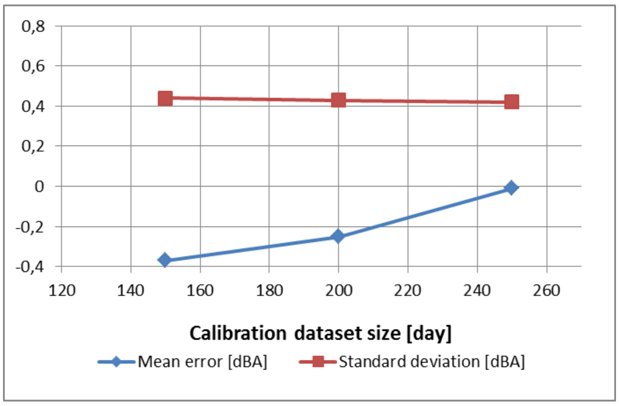


**Fig. 15:** Mean and standard deviation of the error distribution, on the same 50 validation data, as a function of the calibration data size, increasing to the past.

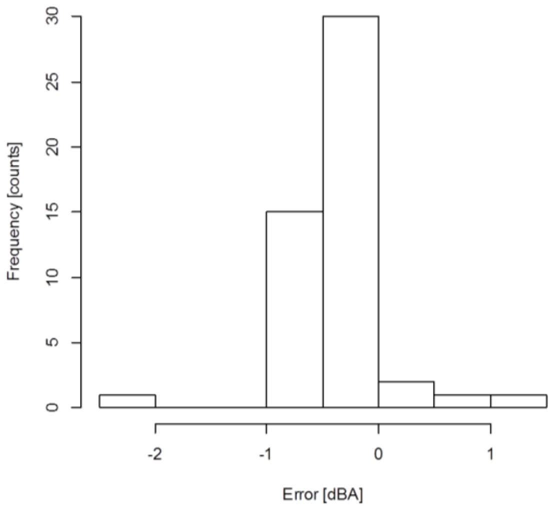


**Fig. 18:** Frequency histogram of the errors calculated on the model validation, obtained calibrating the model on the data from the 101st to the 250th.

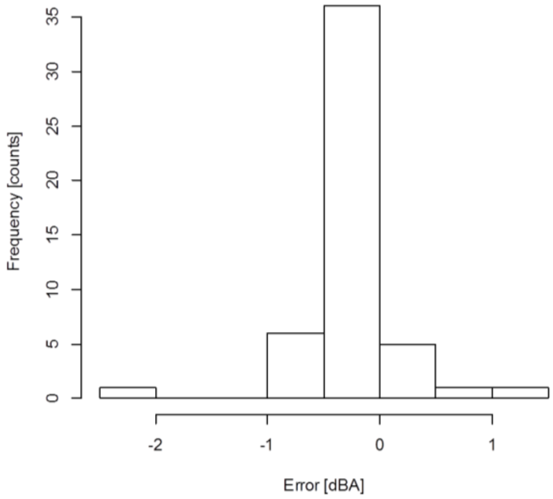


**Fig.19:** Frequency histogram of the errors calculated on the model validation, obtained calibrating the model on the data from the 51st to the 250th.

**Tab.9:** MPE and CVE (error metrics) values, calculated in the calibration and validation phases, for different calibration ranges.

| | Calibration ranges [days] | | |
|---|---|---|---|
| | 101-250 | 51-250 | 1-250 |
| MPE (calibration) | -0.4164 | -0.22847 | -0.02172 |
| MPE (validation) | -0.51539 | -0.34955 | -0.02035 |
| CVE (calibration) | 0.008159 | 0.006848 | 0.006159 |
| CVE (validation) | 0.007955 | 0.006866 | 0.005691 |

Observing the curves in Figures 15, 16 and 9 is evident that the model presented in this paper gives the best forecasts increasing the calibration dataset size. This is due to the fact that the behaviour of the noise levels does not drastically change in the calibration datasets, both in terms of trend and periodicity. This result is confirmed by the error analysis reported in Table 8: mean and standard deviation of the error decrease, in absolute value, if there is the possibility to calibrate the model parameters on a larger dataset (Fig. 17).

Frequency histograms of the errors evaluated in the validation phase (on the data from the 251st to the 300th period), reported in Figures 18 and 19, show that the assumption of normal error distribution is reasonable.

Model performances, evaluated in accordance with error metrics described in section 2, are shown in Table 9. The error dispersion highlighted by CVE decreases constantly, when increasing the calibration dataset size. The MPE metric slightly reduces (in absolute value) in the first step, i.e. adding 50 data to the calibration dataset. In the second step, when 100 data are added to the calibration dataset, MPE drastically approaches zero. This result is obtained because the data inserted in the last step of calibration (from the 1st to the 50th) have a behaviour very similar to the data in the validation period.

# 4 Conclusions

In this paper the authors presented a model based on Time Series Analysis (TSA), founded on a mixed approach, multiplicative between trend and seasonality and additive for the error (irregularity), and its application to an acoustical noise levels dataset.

The analysed data had an almost stationary trend, with a slightly increasing regression line. The model tuned on the first 150 periods provided a good prediction on the next 171 data, with an average error of 0.33 dBA and a standard deviation of 0.52 dBA. In addition, the error histogram shape and kurtosis and skewness indexes, confirmed the hypothesis of normal distribution of errors.

An analysis of the model performances varying the calibration data set size has been performed and reported. An improvement of the error statistics was expected and confirmed in the last step (largest data set size). Moreover, the growth of spread of actual values in the second validation range, measured by an increase of standard deviation of data, pointed out a degradation of model forecast performance in this step. As expected, the Mean Percentage Error (MPE) on the validation data decreased when the amount of calibration data size grew. The error dispersion (CVE) slightly increased when the model was tuned on 200 data instead of 150, because of the growth of spread of data described above, but significantly reduced when the calibration was carried out on 250 periods.

In addition, a model calibration on three different time ranges, enlarging the calibration dataset to the past, has been performed, validating on the same 50 data. The latter application confirmed the improvement of predictive capabilities of the TSA model when the calibration range increases. This result, that is supported by error distribution statistics and error metrics values, can be explained considering that the behaviour of the noise levels does not present drastic change in the dataset, both in terms of trend and periodicity.

Finally, the proposed model has shown good predictive performances, resulting at the same time easy to be implemented and with a low computational duty. It can be installed and compiled in low performance computers and laptops, giving the opportunity of being implemented in “on site” monitoring and forecasting stations, able to transmit measured data and expected values for future periods. In this case, the first calibration data range variation analysis may help the operator to understand when the model will begin to give reliable forecasts, starting from the installation and turn on date, and which range is more appropriate to be considered in the calibration phase. Moreover, the second analysis on variable calibration dataset, may be useful to understand how many past data have to be taken into account to have the best prediction on a given future interval.

# Acknowledgment

The authors are grateful to the local government of Messina, for having made available the long term noise levels measured in the city. A shortened version of this work ([42]) was presented at the 2nd Int. Conf. on Acoustics, Speech and Audio Processing (ASAP '14), Salerno, Italy, June 3-5, 2014.

*References:*

[1] *Demographia World Urban Areas*, 9th Annual Edition, 2013.

[2] Quartieri J., Troisi A., Guarnaccia C., D’Agostino P., D’Ambrosio S., Iannone G., *Development of an Environmental Quality Index Related to Polluting Agents*, Proceedings of the WSEAS International Conference on “Environment, Ecosystem and Development” (EED’09), Puerto de la Cruz, Tenerife (Spain), 14-16 December 2009, pp. 153-161.

[3] Guarnaccia C., Mastorakis N. E., Quartieri J., Noise Sources Analysis in a Wood Manufacturing Company, *International Journal of*

*Mechanics*, Issue 2, Vol. 7, pp 37-44 (2013), ISSN: 1998-4448.

[4] Quartieri J., Mastorakis N. E., Iannone G., Guarnaccia C., D'Ambrosio S., Troisi A., Lenza T.L.L, *A Review of Traffic Noise Predictive Models*, Proceedings of the 5th WSEAS International Conference on "Applied and Theoretical Mechanics" (MECHANICS'09), Puerto de la Cruz, Tenerife, Spain, 14-16 December 2009, pp. 72-80.

[5] Guarnaccia C., Lenza T.LL., Mastorakis N.E., Quartieri J., A Comparison between Traffic Noise Experimental Data and Predictive Models Results, *International Journal of Mechanics*, Issue 4, Vol. 5, pp. 379-386 (2011), ISSN: 1998-4448.

[6] Guarnaccia C., Advanced Tools for Traffic Noise Modelling and Prediction, *WSEAS Transactions on Systems*, Issue 2, Vol.12, 2013, pp. 121-130.

[7] Iannone G., Guarnaccia C., Quartieri J., Speed Distribution Influence in Road Traffic Noise Prediction, *Environmental Engineering And Management Journal*, Vol. 12, Issue 3, 2013, pp. 493-501.

[8] Quartieri J., Iannone G., Guarnaccia C., *On the Improvement of Statistical Traffic Noise Prediction Tools*, Proceedings of the 11th WSEAS International Conference on "Acoustics & Music: Theory & Applications" (AMTA '10), Iasi, Romania, 13-15 June 2010, pp. 201-207.

[9] Guarnaccia C., Analysis of Traffic Noise in a Road Intersection Configuration, *WSEAS Transactions on Systems*, Issue 8, Volume 9, (2010), pp.865-874, ISSN: 1109-2777.

[10] Quartieri J., Mastorakis N. E., Guarnaccia C., Troisi A., D'Ambrosio S., Iannone G., Traffic Noise Impact in Road Intersections, *International Journal of Energy and Environment*, Issue 1, Volume 4 (2010), pp. 1-8.

[11] Iannone G., Guarnaccia C., Quartieri J., *Noise Fundamental Diagram deduced by Traffic Dynamics*, in "Recent Researches in Geography, Geology, Energy, Environment and Biomedicine", Proceedings of the 4th WSEAS Int. Conf. on Engineering Mechanics, Structures, Engineering Geology (EMESEG '11), Corfù Island, Greece, July 14-16, 2011, pp. 501-507.

[12] Quartieri J., Mastorakis N.E., Guarnaccia C., Iannone G., *Cellular Automata Application to Traffic Noise Control*, Proc. of the 12th Int. Conf. on "Automatic Control, Modelling & Simulation" (ACMOS '10), Catania (Italy), 29-31 May 2010, pp. 299-304.

[13] Box, G. E. P., and Jenkins, G., *Time Series Analysis: Forecasting and Control*, Holden-Day, 1976.

[14] Chatfield, C., *The Analysis of Time Series: an Introduction*, Chapman & Hall, New York, 1975.

[15] Di Matteo T., Aste T., Dacorogna M.M., Scaling behaviors in differently developed markets, *Physica A: Statistical Mechanics and its Applications*, Vol. 324, Issues 1–2, 2003, pp. 183-188.

[16] Milanato D., *Demand Planning. Processi, metodologie e modelli matematici per la gestione della domanda commerciale*, Springer, Milano, 2008, in Italian.

[17] Chase R. B., Aquilano N. J., *Operations Management for Competitive Advantage*, Irwin Professional Pub, 10th edition, 2004.

[18] Montgomery D. C., Johnson L. A., Gardiner J. S., *Forecasting and Time Series Analysis*, Mcgraw-Hill, 1990.

[19] http://mobilitamessina.it/index.php/monitoraggio-ambientale

[20] *Directive 2002/49/EC of the European Parliament and of Council of June 25 2002 relating to the assessment and management of environmental noise*, Official Journal of the European Communities, L189/12-25,18.7.2002

[21] D.M. 16.03.1998, *Tecniche di rilevamento e di misurazione dell'inquinamento acustico*, Italian Ministry decree, 1998, in italian.

[22] Guarnaccia C., *Acoustical Noise Analysis in Road Intersections: a Case Study*, Proceedings of the 11th WSEAS International Conference on "Acoustics & Music: Theory & Applications" (AMTA '10), Iasi, Romania, 13-15 June 2010, pp. 208-215.

[23] Guarnaccia C., *New Perspectives in Road Traffic Noise Prediction*, in "Latest advances in Acoustics and Music", proceedings of the 13th Int. Conf. on Acoustics & Music: Theory & Applications (AMTA '12), Iasi, Romania, 13-15 June 2012. ISBN: 978-1-61804-096-1, pp. 255-260

[24] Quartieri J., Mastorakis N.E., Guarnaccia C., Iannone G., *Church Acoustics Measurements and Analysis*, Proceedings of the 11th WSEAS International Conference on "Acoustics and Music: Theory and Applications" (AMTA'10), Iasi, (Romania), 13-15 June 2010, pp: 216-224.

[25] Quartieri J., Guarnaccia C., D'Ambrosio S., Iannone G., *Room Acoustics Experimental Study: Characterization of the Sound Quality in a New Built Church*, Proceedings of the 10th WSEAS International Conference on "Acoustics & Music: Theory & Applications" (AMTA '09), Prague (Rep.Ceca), 23-25 March 2009, pp 108-115.

[26] Quartieri J., D'Ambrosio S., Guarnaccia C., Iannone G., Experiments in Room Acoustics: Modelling of a Church Sound Field and

Reverberation Time Measurements, *WSEAS Transactions on Signal Processing*, Issue 3, Vol.5, pp. 126-135 (2009).

[27] Quartieri J., Troisi A., Guarnaccia C., Lenza TLL, D'Agostino P., D'Ambrosio S., Iannone G., *Analysis of Noise Emissions by Train in Proximity of a Railway Station*, Proceedings of the 10th International Conference on "Acoustics & Music: Theory & Applications" (AMTA '09), Prague (Rep.Ceca), 23-25 March 2009, pp: 100-107.

[28] Quartieri J., Troisi A., Guarnaccia C., Lenza TLL, D'Agostino P., D'Ambrosio S., Iannone G., *An Italian High Speed Train Noise Analysis in an Open Country Environment*, Proceedings of the 10th International Conference on "Acoustics & Music: Theory & Applications" (AMTA '09), Prague (Rep.Ceca), 23-25 March 2009, pp: 92-99.

[29] Quartieri J., Troisi A., Guarnaccia C., Lenza TLL, D'Agostino P., D'Ambrosio S., Iannone G., An Acoustical Study of High Speed Train Transits, *WSEAS Transactions on Systems*, Issue 4, Vol.8, pp. 481-490 (2009).

[30] Quartieri J., Troisi A., Guarnaccia C., Lenza TLL, D'Agostino P., D'Ambrosio S., Iannone G., Application of a Predictive Acoustical Software for Modelling Low Speed Train Noise in an Urban Environment, *WSEAS Transactions on Systems*, Issue 6, Vol.8, pp. 673-682 (2009).

[31] Quartieri J., Mastorakis N. E., Guarnaccia C., Troisi A., D'Ambrosio S., Iannone G., *Road Intersections Noise Impact on Urban Environment Quality*, Proceedings of the 5th WSEAS International Conference on "Applied and Theoretical Mechanics" (MECHANICS '09), Puerto de la Cruz, Tenerife, Spain, 14-16 December 2009, pp. 162-171.

[32] Quartieri J., Sirignano L., Guarnaccia C., *Equivalence between Linear and Curved Sources in Newtonian Fields: Acoustics Applications*, Proc. Of the Int. Conf. on Engineering Mechanics, Structures, Engineering Geology (EMESEG '08), Heraklion, Crete Island, Greece, July 22-24, 2008, pp: 393-395.

[33] Quartieri J., Guida M., Guarnaccia C., D'Ambrosio S., Guadagnuolo D., *Complex Network Applications to the Infrastructure Systems: the Italian Airport Network case*, Proc. of the Int. Conf. on Urban Planning and Transportation (UPT'07), Heraklion, Crete Island, Greece, July 22-24, 2008, pp: 96-100.

[34] Quartieri J., Mastorakis N.E., Iannone G., Guarnaccia C., *A Cellular Automata Model for Fire Spreading Prediction*, in "Latest Trends on Urban Planning and Transportation", Proc. of the 3rd Int. Conf. on "Urban Planning and Transportation" (UPT '10), Corfù, Greece, 22-24 July 2010. ISBN: 978-960-474-204-2/ISSN: 1792-4286, pp. 173-179.

[35] Iannone G., Troisi A., Guarnaccia C., D'Agostino P. P., Quartieri J., An Urban Growth Model Based on a Cellular Automata Phenomenological Framework, *Int. Journal of Modern Physics C*, Volume 22, Issue 05, pp. 543-561 (2011). DOI: 10.1142/S0129183111016427.

[36] Quartieri J., Guida M., Guarnaccia C., D'Ambrosio S., Guadagnuolo D., Topological Properties of the Italian Airport Network studied via Multiple Addendials and Graph Theory, *International Journal of Mathematical Models and Methods in Applied Sciences*, Issue 2, Vol.2, pp 312-316 (2008).

[37] Guarnaccia C., Mastorakis N.E., Quartieri J., Wind Turbine Noise: Theoretical and Experimental Study, *International Journal of Mechanics*, Issue 3, Vol.5, pp. 129-137 (2011).

[38] Quartieri J., Sirignano L., Guarnaccia C., Infinitesimal Equivalence between Linear and Curved Sources in Newtonian Fields: Application to Acoustics, *International Journal of Mechanics*, Issue 4, Vol.1, pp. 89-91 (2007) , ISSN: 1998-4448.

[39] D'Ambrosio S., Guarnaccia C., Guida D., Lenza T.L.L., Quartieri J., System Parameters Identification in a General Class of Non-linear Mechanical Systems, *International Journal of Mechanics*, Issue 4, Vol. 1, pp 76-79 (2007).

[40] Guarnaccia C., Mastorakis N. E., Quartieri J., *A Mathematical Approach for Wind Turbine Noise Propagation*, in Applications of Mathematics and Computer Engineering, American Conference of Applied Mathematics (AMERICAN-MATH '11), Puerto Morelos, Mexico, 29-31 January 2011, pp. 187-194.

[41] Guarnaccia C., Quartieri J., Rodrigues E. R. and Tepedino C., *Time Series Model Application to Multiple Seasonality Acoustical Noise Levels Data Set*, in "Latest Trends in Circuits, Systems, Signal Processing and Automatic Control", proceedings of the 2nd Int. Conf. on Acoustics, Speech and Audio Processing (ASAP '14), Salerno, Italy, June 3-5, 2014 , ISSN: 1790-5117, ISBN: 978-960-474-374-2, pp. 171-180.

[42] Guarnaccia C., Quartieri J., Mastorakis N. E. and Tepedino C., *Acoustic Noise Levels Predictive Model Based on Time Series Analysis*, in "Latest Trends in Circuits, Systems, Signal Processing and Automatic Control", proceedings of the 2nd Int. Conf. on Acoustics, Speech and Audio Processing (ASAP '14), Salerno, Italy, June 3-5, 2014 , ISSN: 1790-5117, ISBN: 978-960-474-374-2, pp. 140-147.

[43] Guarnaccia C., Quartieri J., Mastorakis N. E., *Comparison of Acoustic Barriers Noise Reductions Evaluated by Different Calculation Methods*, in “Latest Trends on Systems”, Proc. of the 18th Int. Conf. on Circuits, Systems, Communications and Computers (CSCC'14), Santorini, Greece, 17-21 July 2014, pp. 443-449.

[44] Guarnaccia C., Quartieri J., Cerón Bretón J. G., Tepedino C., Cerón Bretón R. M., *Time Series Predictive Model Application to Air Pollution Assessment*, in “Latest Trends on Systems”, Proc. of the 18th Int. Conf. on Circuits, Systems, Communications and Computers (CSCC'14), Santorini, Greece, 17-21 July 2014, pp. 499-505.