Modelling of ambient noise levels in urban environment

Shashi Kant Tiwari, L.A Kumaraswamidhas and Naveen Garg

December 5, 2019

Modelling of ambient noise levels in urban environment

S.K.Tiwari[1], L.A. Kumaraswamidhas[1,*], N. Garg[2]
[1]Indian Institute of Technology (ISM), Dhanbad-826 001
[2] CSIR-National Physical Laboratory, New Delhi-110 012.

[*]email: shashikanttiwari61@gmail.com

The study conducts time-series approach for analysing one year noise monitoring data. Support Vector Machine (*SVM*) technique is used as a modelling technique for time-series approach. The noise data is trained using 10-fold cross validation to get optimum hyper-parameters (γ, ε, C). The performance and accuracy of model is determined by statistical parameters like *MSE, RMSE, MAPE in %, R^2*. The paper predicts an error of ± 2dB(A) with the implementation of Support Vector Machine (*SVM*).

Keywords: Noise Monitoring, Support Vector Machine, Ambient Noise Levels

1. Introduction

Noise pollution has significantly increased especially in the urban areas in Indian scenario. With advancement of vehicles and urbanization in cities, there has been a quick addition in traffic volume. Regardless of the way that transportation is an essential part of urban society, its superiority is obscured by its negativity. Inappropriate placement of vehicles at different locations nearby roads is one of the major cause of traffic jam. Some studies affirm that noise pollution has adverse influence on human health [1, 2]. It comprises of slant stress impact, sleeping disturbances which clearly cause 'prompt effect' on mental and physical perspective. The Central Pollution Control Board has directed noise levels for different zones i.e Silence, Industrial, Commercial, Residential zones and carried many studies for noise monitoring in Indian scenario [3]. Garg et al. [4] discussed about the pilot project on the establishment of National Ambient Noise Monitoring Network (*NANMN*) at 35 locations across the 7 major cities of the country. The European Environmental Noise Directive 2002/49/EC [5] gives direction for noise mapping that include future plans with financial information and cost-effective assessment. European Directive permits to survey and to look at, inside EU Member States, noise exposure data, particularly for the future execution steps, when noise maps should be drawn up with the basic evaluation strategies.

There are various techniques used for Noise assessment and monitoring. Some uses long-term noise monitoring strategy while Garg et al. [6] emphasised on short-term noise monitoring strategy as a reliable strategy within an accuracy of ±2 dB(A). The high costs of installing and maintaining permanent networks is primarily the main reason for analysing the suitability of short-term strategies to ascertain whether they can provide a suitable and reliable alternative or not as compared to the long-term noise monitoring. There are some illustrations whereby extensive networks have been installed (Czyzewski et al. [7]). Morillas and Gajardo [8] evaluated 90% probability interval for random 9 days data to measure L_{den}. Hence, there is a need of alternative approach to predict and forecast ambient noise level by using time-series approach. DeVor et al. [9] used *ARMA* (auto regressive moving average) and Garg et al. [10] used Autoregressive Integrated Moving Average (*ARIMA*) model to predict and forecast noise level in time series analysis. This study explore Support Vector Machine (*SVM*) technique for forecasting and determining the accuracy and performance of the predicted noise level.

2. Methodology

Support Vector Machine: The idea of support vector machine (*SVM*) is mapping a non-linear data set. The approach focuses to solve a regression using linear function. The hyper-plane also known as classifier separates classes to get an optimal solution. The data is spitted into training and testing data. Suppose X_i represent input data (i= 1,……,n) where *n* is the number of training data points. The function of hyper-plane is as [11]:

$$Y(x) = w^T X_i + b \tag{1}$$

Where, *w* is the orientation and *b* is the position of hyper-plane classifying the training data into two classes. C_1 is the positive class and C_2 is the negative class. The main focus of Support Vector Machine is to find a new classifier.

$$Y(x_1) = w^T x_i + b > 0 \tag{2}$$

$x_1 \in C_1$, if x1 lies on the positive side of the hyper-plane.

$$Y(x_2) = w^T x_i + b < 0 \quad (3)$$

$x_2 \in C_2$, if x1 lies on the positive side of the hyper-plane.

The dot product of two input $\psi(x)$ and $\psi(y)$ vectors is Kernel Function that should affirms Mercer's condition. Mainly four kernels are utilized in support vector machine (SVM) modelling which are as follows [12]:

Linear Kernel: $K(x,y) = x^T y$ (4)

Radial Basis Function Kernel: $K(x,y) = \exp(-\gamma\| x - y^2)$, $\gamma > 0$ (5)

Polynomial Kernel: $K(x,y) = (\gamma x^T y + r)^d$, $\gamma > 0$ (6)

Sigmoid Kernel: $K(x,y) = \tanh(\gamma x^T y + r)$ (7)

Here, the kernel parameters are d, r and γ. Kernel parameters have important significance in the performance of support vector machine model. The complexity of best parameter is controlled by the kernel functions.

3. Results and Discussion

Fig 1(a) shows variation plot of L_{day} in dB(A) for 365 days for a Residential site in Delhi. Maximum value of L_{day} is 72 dB(A) and minimum value is 60 dB(A). While, Fig 1(b) shows variation plot of L_{night} in dB(A) for 365 days for a Residential site in Delhi. Maximum value of L_{night} is 72 dB(A) and minimum value is 60 dB(A).

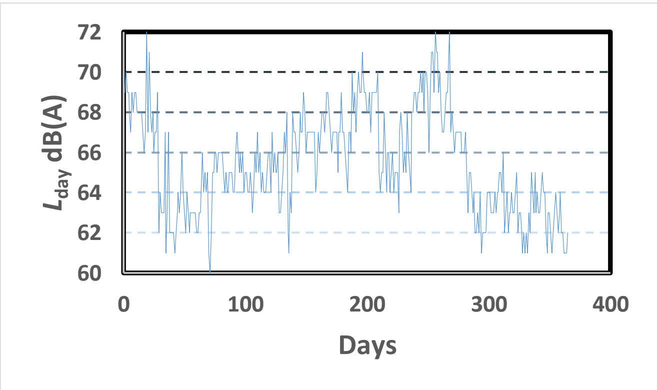

Fig. 1(a): Time sequence plot of L_{day} in dB(A) for 365 days for a Residential site in Delhi

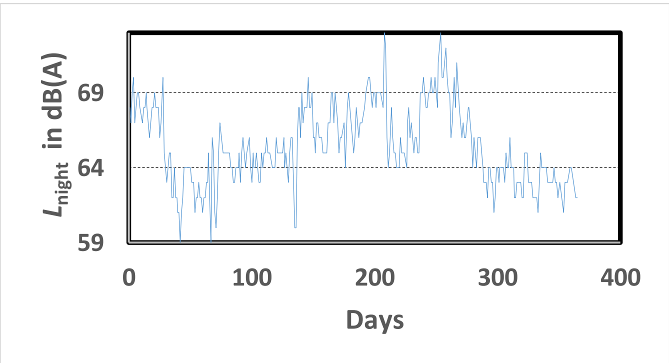

Fig. 1(b): Time sequence plot of L_{night} in dB(A) for 365 days for a Residential site in Delhi

The Kernel applied in the study is Radial Basis Function (RBF) kernel. It has better performance in comparison to the others kernel functions. In RBF kernel three hyper-parameters has been used to analyse the performance of SVM model. These three hyper-parameters are: Gamma (γ), Epsilon (ε) and Cost (C). The first stage is to find an optimized parametric combination of these hyper-parameters. Hit and Trial approach was attempted to get the optimum value of hyper-parameters. The parametric combination (γ, ε, C) of optimized hyper-parameters is (2^5, 0.4,2^5) for both day and night. Fig.2(a)- 2(b) shows the plot of predicted day and night ambient noise level in comparison to observed noise levels.

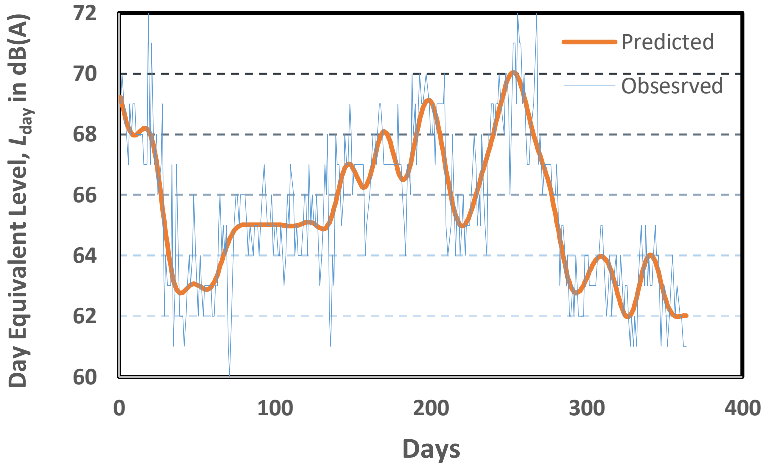

Fig. 2(a): Comparison of Measured (blue line) and Predicted (red line) values of L_{day} in dB(A)

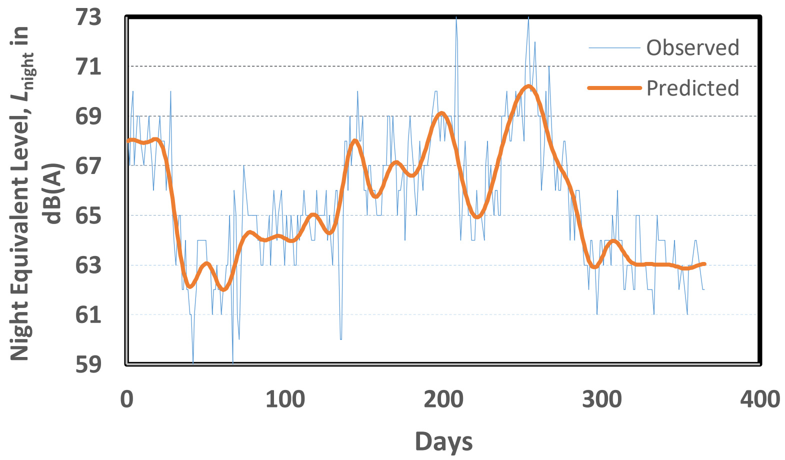

Fig. 2(b): Comparison of Measured (blue line) and Predicted (red line) values of L_{night} in dB(A)

The input data is divided into testing and training data. 90% of the input data is taken as training data and 10% of the data is used as testing data. Mean square error (*MSE*), Root mean square error (*RMSE*), Mean average percentage error (*MAPE in%*), Coefficient of determination (R^2) are the parameters that ascertain the efficiency of the model.

Table 1: *SVM* Model Statistics for Training data

Statistical Parameter	L_{day} dB(A)	L_{night} dB(A)
MSE in dB(A)^2	1.640	1.664
RMSE in dB(A)	1.281	1.290
MAPE in %	1.47	1.43
Maximum Error in dB(A)	4.54	5.37
Minimum Error in dB(A)	-4.37	-5.96
R^2	0.6	0.4

Table.1 shows the statistics performance of training data for both day and night noise level. The maximum error is 4.54 dB(A) for day and 5.37 dB(A) for night but the MSE and RMSE error lies within ± 2 dB(A) which is sought of reliable accuracy. The probability of the training data determined from input data can be taken as per the study analysis. There is no particular approach to determine the probability like in the present study the training data is taken 90%. The determination of co-efficient is 0.6 for day and 0.4 for night which implies better performance of the classifier.

Table 2: *SVM* Model Statistics for Testing data

Statistical parameter	L_{day} dB(A)	L_{night} dB(A)
MSE in $dB(A)^2$	2.316	1.695
RMSE in dB(A)	1.522	1.302
MAPE in %	1.45	1.82
Maximum Error in dB(A)	2.47	1.97
Minimum Error in dB(A)	-3.20	-3.09
R^2	0.6	0.5

To test the model, the testing data is taken 10% of the input data. The SVM model predicts an error of ± 2 dB(A) for testing data as well. The determination of co-efficient is 0.6 for day and 0.5 for night which implies better performance of the classifier of predicted testing data. Hence, it can be observed that the support vector machine is a good approach for forecasting of ambient noise level within an accuracy of ± 2 dB(A).

4. Conclusion

In the study, Support Vector Machine (*SVM*) is used as a time-series modelling technique for the statistical analysis of one year noise monitoring data set. *SVM* is an outperforming technique that can profoundly predict the ambient noise levels L_{day} and L_{night}. The parametric combination for both day and night is (2^5, 0.4,2^5). Meanwhile, this is the best set of hyper-parameter for classifier represent a similar trend as the observed pattern. The application of Support Vector Machine (*SVM*) has rarely been used in the determination of ambient noise level. The result shows that this model can be used as a better fitting model for predicting and forecasting noise levels. The work also emphasis on the use of Radial Basis Function (RBF) kernel to analyse *SVM* model. The performance of model is determined by the statistical parameters like *MSE, RMSE, MAPE in % and* R^2.

References

1. Ising, H., & Kruppa, B. (2004). Health effects caused by noise: evidence in the literature from the past 25 years. Noise and Health, 6(22), 5.

2. World Health Organization. Burden of disease from environmental noise. Geneva: WHO; 2011.

3. Central Pollution Control Board, Annual Report, 2011–2012. http://cpcb.nic.in/upload/AnnualReports/AnnualReport 43 AR 2011-12 English.pdf [access on 02.03.2016].

4. Garg, N., Sinha, A. K., Gandhi, V., Bhardwaj, R. M., & Akolkar, A. B. (2016). A pilot study on the establishment of national ambient noise monitoring network across the major cities of India. Applied Acoustics, 103, 20-29.

5. European Noise Directive. Assessment and Management of Environmental Noise. 2002/49/EU, Official Journal of European Communities, (2002). DIRECTIVE 2002/49/EC of the European parliament and of the council of 25 June 2002 relating to the assessment and management of environmental noise.

6. N. Garg, T.K. Saxena and S. Maji. Long- term versus short- term noise monitoring: Strategies and implications in India. Noise Control Engineering Journal, Volume 63, Number 1, 1 January 2015, pp. 26-35(10).

7. A. Czyzewski., J. Kotus, M. Szczodrak. On-line urban acoustic noise monitoring system. Noise Control Eng. J. 60 (1), 69–84, (2012).

8. Morillas, J. B., & Gajardo, C. P. (2014). Uncertainty evaluation of continuous noise sampling. Applied Acoustics, 75, 27-36.

9. DeVor, R. E., Schomer, P. D., Kline, W. A., & Neathamer, R. D. (1979). Development of temporal sampling strategies for monitoring noise. The Journal of the Acoustical Society of America, 66(3), 763-771.

10. Garg, N., Soni, K., Saxena, T. K., & Maji, S. (2015). Applications of Autoregressive integrated moving average (ARIMA) approach in time-series prediction of traffic noise pollution. Noise Control Engineering Journal, 63(2), 182-194.

11. Samsudin, R., Shabri, A., & Saad, P. (2010). A comparison of time- series forecasting using support vector machine and artificial neural network model. Journal of applied sciences, 10(11), 950-958.

12. Ding, Y., Song, X., & Zen, Y. (2008). Forecasting financial condition of Chinese listed companies based on support vector machine. Expert Systems with Applications, 34(4), 3081-3089.