Application of selected computational intelligence methods to sound level modelling based on traffic intensity in thoroughfare

Michał Kekez[1,*]

[1]Kielce University of Technology, Faculty of Mechatronics and Mechanical Engineering, Aleja Tysiąclecia Państwa Polskiego 7, 25314 Kielce, Poland

Abstract. The aim of the paper was to build the models of sound pressure level as a function of traffic intensity in thoroughfare. The models were built by using artificial analytical models or regression trees. The former included Nordic Prediction Method. The latter were represented by Random Forest and Cubist. The analysis of accuracy of all obtained models was conducted. The best models can be used in the process of reconstruction of equivalent sound level data.

Keywords: traffic noise, nordic prediction method, random forest

1 Traffic intensity and sound level measurement results

The level of road traffic noise depends on many parameters. Some of them are constant during long period of time (road geometry, type and condition of road surface), but other parameters, which are connected with traffic intensity, can change very fast: number of trucks, number of other vehicles, and average vehicle speed.

In these paper the measurement data recorded at Krakowska Street in Kielce (Poland) were used for building and evaluation of the models of sound pressure level as a function of traffic intensity in thoroughfare. The number of vehicles, their category, speed, as well as equivalent sound pressure level values were measured and recorded in a monitoring station comprising a sound level meter (SVAN 958A, digital, four-channel, class-1, vibration and sound meter), a weather station, and road radar [2]. A-weighted equivalent sound level pressure values $L_{Aeq,i}$ were recorded every minute. Next, by using equation (1)

$$L_{Aeq} = 10 \cdot \log\left(\frac{1}{N}\sum_{i=1}^{N} 10^{0.1L_{Aeq,i}}\right) \quad (1)$$

L_{Aeq} values were calculated for longer periods of time, namely for the following 24h sub-intervals: day (6-18), evening (18-22) and night (22-6). For the same time periods, vehicle speed values were averaged, and numbers of vehicles were summed up.

* Corresponding author: m.kekez@tu.kielce.pl
Reviewers: *Maciej* Berdychowski, *Ireneusz* Malujda

During the year 2013, for day sub-intervals, there were only 296 complete records, i.e. records (days) without missing measurement values, as shown in Fig. 1. Each record contains values of: A-weighted equivalent sound level, number of light vehicles, number of heavy vehicles, and average speed of all vehicles. In incomplete records, the most frequently missing value is L_{Aeq}, while the traffic intensity parameters (average speed, number of light and heavy vehicles) are available for almost all days in 2013. Finding the model which describes relationship between sound level and traffic intensity parameters can benefit in the process of reconstruction of missing equivalent sound level data.

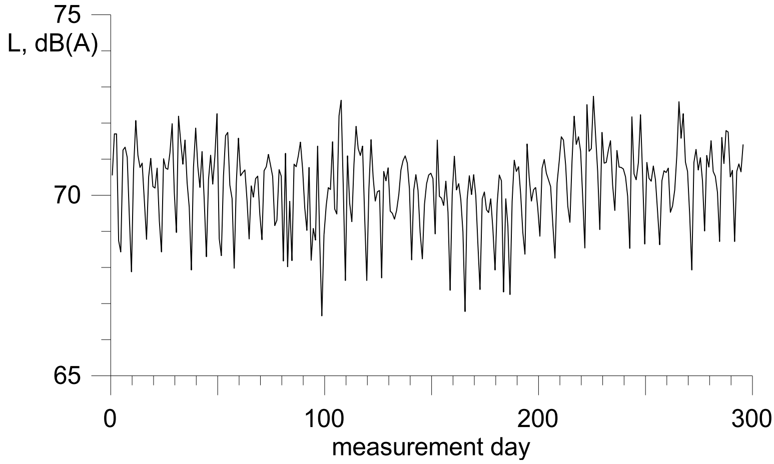

Fig. 1. L_{Aeq} calculated from measurements for day sub-intervals (6-18) in year 2013

The distribution of values from the above mentioned 296 complete records is shown in Fig. 2. The L_{Aeq} varies from 66.7 to 72.7 dB. Average speed of vehicles belongs mainly to the range 50.0-65.2 km/h, but there are 6 days with much lower values, from 32.2 to 40.3 km/h. The number of heavy vehicles predominantly belongs to range 140 to 1340, but there are 6 days with higher values (from about 2000 to 2690). The same 6 days have the lowest vehicle average speed values (below 40.3 km/h). The number of light vehicles changes from about 8780 to 28460.

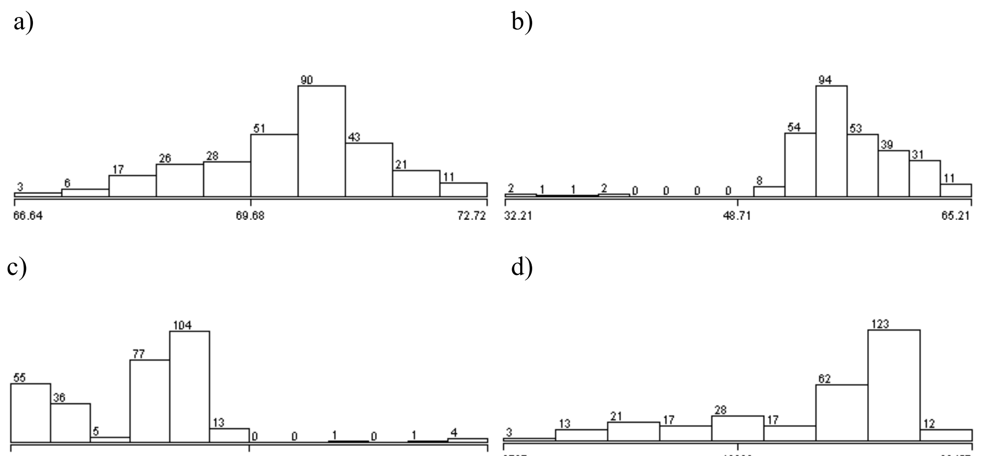

Fig. 2. Distribution of values for day sub-interval in 296 days in year 2013: a) equivalent sound pressure level A, L_{Aeq}, b) average speed of vehicles, c) number of heavy vehicles, d) number of light vehicles

Selected plot matrices are shown in Fig. 3. For average speed values above 50 km/h, L_{Aeq} decreases with increase of average vehicle speed, which may be counter-intuitive, but easy to explain: higher average speed means that traffic density is smaller and less vehicles are passing the monitoring station. Fig. 3b shows that L_{Aeq} increases with increasing number of light vehicles.

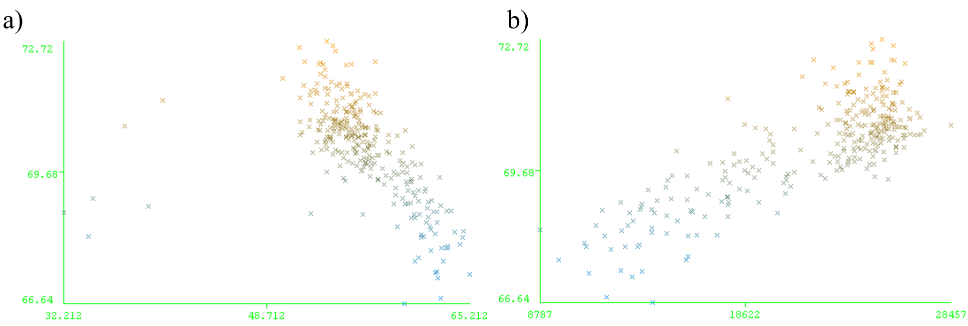

Fig. 3. Plot matrices, vertical axis: L_{Aeq}, horizontal axis: a) average speed of vehicles, b) number of light vehicles

2 Nordic Prediction Method model of sound level

There are some models which allow to calculate equivalent sound level L_{Aeq} on the basis of the parameters related to traffic intensity mentioned in section 1. One of them is Nordic Prediction Method model [1]. It can be expressed (after some transformations) by equations (2)-(6). Let v_1 and v_2 denote speed (in km/h) of heavy and light vehicles, respectively, and q_1 and q_2 – number of heavy and light vehicles per second, respectively. Then, for $v_1 \geq 50$ km/h and $v_2 \geq 40$ km/h:

$$L_1 = 29.5 + 30 \log v_1 + 10 \log q_1 \quad (2)$$

$$L_2 = 31 + 25 \log v_2 + 10 \log q_2 \quad (3)$$

$$L_{Aeq} = 10 \log(10^{0.1L_1} + 10^{0.1L_2}) \quad (4)$$

In case of lower vehicle speeds, $30 \leq v_1 < 50$, and $30 \leq v_2 < 40$, instead of (1) and (2), equations (5) and (6) should be used:

$$L_1 = 80.5 + 10 \log q_1 \quad (5)$$

$$L_2 = 71.1 + 10 \log q_2 \quad (6)$$

When the Nordic Prediction Method model was applied to calculation of L_{Aeq} by using q_1, q_2, and v (assuming $v_1 = v_2$) values taken from each of 296 records mentioned in Section 1, the L_{Aeq} values calculated by the model were on average 2.3 dB higher than recorded values.

Another model, described in [6] requires vehicles speed data separately for each of 4 categories of vehicles, as well as calculations of sound propagation. For these reasons author decided (in Section 3) to propose the new model, inspired by Nordic Prediction Method. Other studies showed that regression as well as artificial neural networks can be used for modelling L_{Aeq} as a function of traffic intensity parameters [7], or even without direct use of vehicle volume and speed [8, 9].

3 Cubist and Random Forest models of sound level

Computational intelligence methods such as artificial neural networks, fuzzy systems, random forest regression, or regression trees can be applied to construction of the model which describes relation between input and output attributes in a given dataset. Two approaches from the field of regression forests and regression trees were employed in this study, namely the Random Forest algorithm implemented in WEKA [3] and Cubist software [4].

Training dataset consisted of 296 records, each of them describing one day sub-interval (6-18) in year 2013. Dataset contains only records for which sound level, number of vehicles, and their speed were recorded (there are no missing values). Each record contained one output attribute, db_A (L_{Aeq}, A-weighted equivalent sound level), and three input attributes, log_N_light, log_N_heavy, log_avg_speed, which are 10-base logarithms of: number of light vehicles, number of heavy vehicles, and average speed of all vehicles, respectively. These attributes were selected according to equations (2)-(4), which suggest that there exists a correlation between these attribute values and L_{Aeq} value.

Several methods were applied to the training set in order to obtain model of L_{Aeq}. First of them was regression tree software Cubist ver. 2.09, which built quite simple model, using only first 200 records from training dataset due to limitations of the free version:

IF log_N_light ≤ 4.327563 THEN
    db_A = 25.68 + 3.25 log_N_heavy + 11.2 log_avg_speed + 3.6 log_N_light
IF log_N_light > 4.327563 THEN
    db_A = 105.277 – 4.73 log_N_heavy – 29.1 log_avg_speed + 6.9 log_N_light

Scatter plot for this model is shown in Fig. 4.

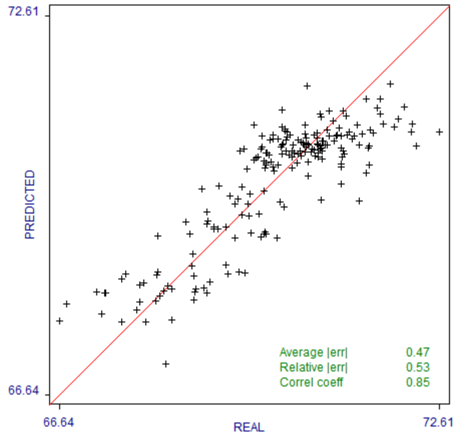

Fig. 4. Scatter plot for model built by Cubist 2.09 for the first 200 records

Next model, built by Cubist 2.07 using 296 records from training dataset, was even simpler:

db_A = 31.157 + 9 log_N_light

Other models, created by Random Forest method [10], also using 296 records from training dataset, were much more complicated (100 large regression trees i.e. one ‘random forest’ in each model). Random Forest method was run with random selection of k = 0, 1, 2, and 3 attributes. Accuracy of all elaborated models, expressed by mean absolute error (MAE), root mean square error (RMSE), and correlation coefficient, is shown in Table 1.

Table 1. Accuracy of regression trees models.

Model	Test,10-fold cross validation			Training data		
	MAE [dB]	RMSE [dB]	correlation coefficient	MAE [dB]	RMSE [dB]	correlation coefficient
Cubist 2.09 (first 200 records)	0.53	0.67	0.81	0.47	0.60	0.85
Cubist 2.07 (all 296 records)	0.54	0.69	0.79	0.52	0.67	0.81
RandomForest k: 0, 2	0.52	0.67	0.81	0.19	0.25	0.98
RandomForest k: 1	0.51	0.65	0.82	0.19	0.24	0.98
RandomForest k: 3	0.53	0.68	0.80	0.19	0.25	0.98

Values of L_{Aeq}, measured, and predicted by two selected models, are shown in Fig. 5.

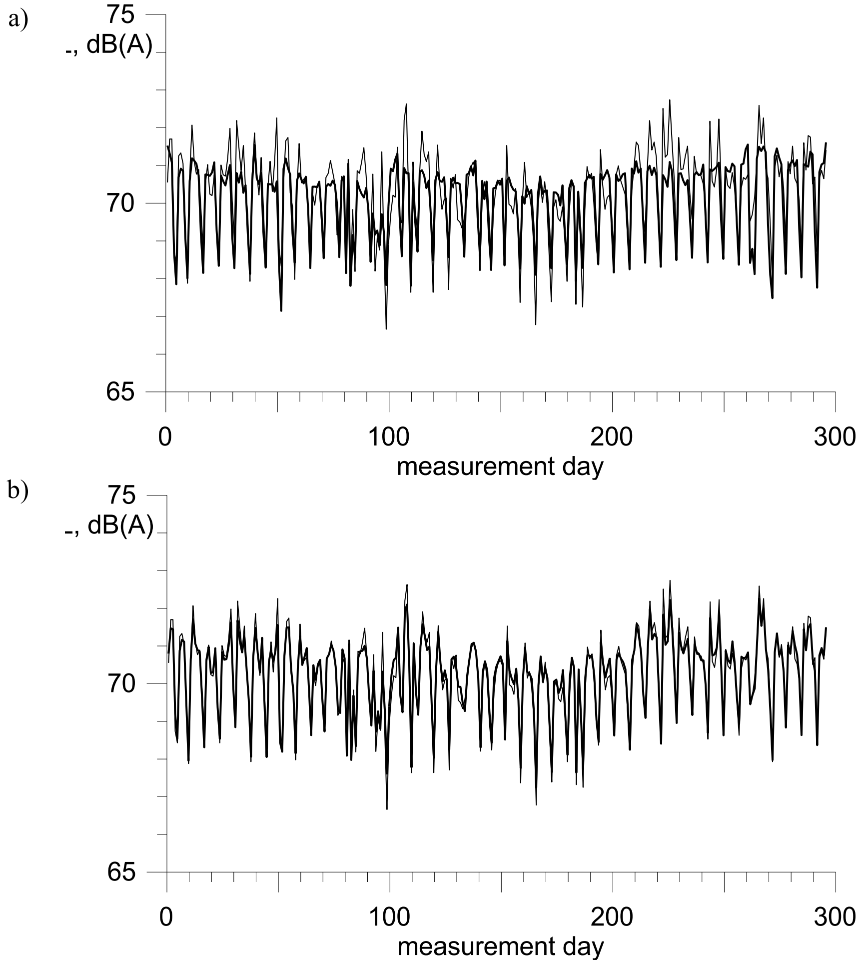

Fig. 5. L_{Aeq} values for day sub-interval in 296 days in year 2013: measured (thin line), and calculated by the models (thick line): Cubist 2.09 model (a), and Random Forest model with random selection of k = 1 attributes (b)

Analysis of Fig. 5 and Table 1 shows that Cubist 2.09 model is quite accurate on its training data (first 200 records) and test data (next 96 records). Random Forest model with random selection of $k = 1$ attributes is over fitted to training data (MAE equals 0.19 dB), but its accuracy on the test set (obtained by 10-fold cross validation) is much lower (MAE equal to 0.51 dB) but still slightly better than Cubist's.

4 Conclusions

All models produced by Cubist and Random Forest are very accurate (mean absolute error below 0.6 dB and RMSE below 0.7 dB on the test set) and can be used in the process of reconstruction of missing equivalent sound level data. The models were elaborated for given thoroughfare and time of day, but the same methodology can be applied for other 24h sub-intervals and other roads.

References

1. H. L. Nielsen, *Road traffic noise: Nordic prediction method* (1997) ISBN 10:9291208361
2. A. Bąkowski, L. Radziszewski, Z. Skrobacki, *Assessment of uncertainty in urban traffic noise measurement*. Procedia Engineering, **177**: 281-288 (2017)
3. http://www.cs.waikato.ac.nz/ml/weka/
4. https://www.rulequest.com/cubist-info.html
5. ISO 9612:2009, *Acoustics – Determination of occupational noise exposure – Engineering method*.
6. S. Kephalopoulos, M. Paviotti, F. Anfosso-Lédée, *Common noise assessment methods in Europe (CNOSSOS-EU)*, DOI 10.2788/32029 (print), 10.2788/31776 (online)
7. P. Kumar, S.P. Nigam, N. Kumar, *Vehicular traffic noise modeling using artificial neural network approach*, Transportation Research Part C, **40**: 111–122 (2014)
8. L. Radziszewski, M. Kekez, A. Bąkowski, A. Sapietova, *Modeling of acoustic pressure variability at thoroughfare* (to be published)
9. L. Radziszewski, M. Kekez, A. Sapietova, *Missing values reconstruction in sound level monitoring station by means of intelligent computing*, MATEC Web of Conferences **157**(02042): 1-7 (2018)
10. L. Breiman, *Random Forests*, Machine Learning, **45** (1): 5–32 (2001)