An intelligent hybridized computing techniques for the prediction of roadway traffic noise based on non-linear mutual information

Ibrahim Khalil Umar
Kano State Polytechnic
Vahid Nourani
Near East University: Yakin Dogu Universitesi
Hüseyin Gökçekuş
Near East University: Yakin Dogu Universitesi
SANI ABBA (saniisaabba86@gmail.com)
Yusuf Maitama Sule University Kano: Maitama Sule University Kano https://orcid.org/0000-0001-9356-2798

Research Article

Keywords: Road traffic noise, artificial intelligence, hybrid models, gaussian process regression, Nicosia.

Posted Date: September 27th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-837045/v1

An intelligent hybridized computing techniques for the prediction of roadway traffic noise based on non-linear mutual information

Ibrahim Khalil Umar[1*], Vahid Nourani[1,2], Hüseyin Gökçekuş[1], S.I. Abba[3*]

[1*] Faculty of Civil and Environmental Engineering, Near East University, Near East Boulevard, 99138, Nicosia, North Cyprus
[2] Center of Excellence in Hydro informatics, Faculty of Civil Engineering, University of Tabriz, Tabriz, Iran
[3*]Faculty of Engineering Department of Civil engineering Baze University
*Corresponding author's E-mail:
ibrahimkhalil.umar@neu.edu.tr,sani.abba@bazeuniversity.edu.ng

Abstract

A reliable traffic noise prediction model is one of the decision-making tools used in providing a noise friendly environment. In this study, four linear-nonlinear hybrid models were proposed to capture both linear and nonlinear patterns of the data by summing up the predicted traffic noise from the multilinear regression (MLR) and estimated residuals from four artificial intelligence (AI)-based models. The input variables for the models were volumes of cars, medium vehicles, buses, heavy vehicles, and average speed. Prior to the development of the hybrid model, the potential of Boosted Regression Tree (BRT), Feed Forward Neural Network (FFNN), Gaussian Process Regression (GPR), Support Vector Regression (SVR) and Linear regression models for traffic noise prediction were evaluated and compared with each other. The performances of the single and hybrid models were evaluated using the Nash-Sutcliffe efficiency (NSE), root mean square error (RMSE), mean absolute error (MAE) and relative root mean square error (rRMSE). The results showed that, the hybrid models provide better prediction capability than both the linear and nonlinear models in both calibration and verification stages. MLR-GPR hybrid demonstrated better prediction skill than all other hybrid models with NSE, RMSE, MAE and rRMSE values of 0.9312, 0.0427, 0.0347 and 7.4%, respectively. The study found that, the efficiency of the linear models could be improved up to 27.26% when they are hybridized with the nonlinear models.

Keywords: Road traffic noise, artificial intelligence, hybrid models, gaussian process regression, Nicosia.

1. Introduction

Monitoring, assessing and prediction of the noise pollution which is one of the major environmental pollutions endangering the wellbeing of residents along the major roads of urban cities are essential for management of noise pollution. This can help in reducing the increased health risks such as premature death, cardiovascular diseases, hearing impairment, sleep disturbance (European Environment Agency 2014), tinnitus (Maschke and Widmann, 2013), occasional memory difficulties (Schlittmeier et al. 2015), increased chance of diabetes and annoyance (Sørensen et al. 2013) which are believed to be associated with incessant exposure to the noise pollution. Roadway noise is the main source of the environmental noise pollution in urban areas and is believed to be higher in areas where the total traffic volume as well as the speed are high (Kumar et al. 2014). The three components of noise that make up the traffic noise are noise generated as a result of the interaction of the vehicles' tire with road pavement; aerodynamically generated noise due to the airflow turbulent through and around the vehicle; propulsion noise from engine, exhaust and transmission. Aerodynamically generated noise dominates other forms of the traffic noise in high-speed roads while the tire pavement interaction is the dominant on low-speed roads (Sandberg and Ejsmont, 2002). The physical measurement of the traffic noise can be expensive and time consuming and even impossible along high speed roads (Ahmed and Pradhan 2019). As a result of these problems, researchers have been developing various statistical and regression models aimed at providing cost effective as well as reliable models capable of predicting the traffic noise with higher accuracy.

However, the traffic noise prediction models serve as decision-making tools that help the stakeholders for both decision and policy making for providing friendly road environments. As a result of local conditions such as variations in traffic composition, weather conditions and road geometry from one country to another, the regression models for the prediction of the roadway traffic noise are faced with the non-generalization problem (Hamad et al. 2017). The non-generalization problem of the models has resulted in the development of various mathematical as well as artificial intelligence (AI)-based models for roadway traffic noise prediction in different regions and countries. Traffic volume, traffic composition, vehicular speed, distance from the noise source, reflective surface, temperature, building facade, gradient, honking and acceleration/deceleration or combination of some of the parameters are the major inputs parameters considered for the development of both statistical and regression models for roadway traffic noise prediction (Gundogdu et al. 2005). Other variables such as driving behavior, driver's skills, vehicle maintenance duties, speed limits and road geometry may also have effects on the mentioned parameters (Covaciu et al. 2015).

The most commonly used models for estimation of roadway traffic noise in the literature includes the German RLS 90, French NMPB-Routes 96 which was approved to be used in many European countries by the environmental noise legislation of the European Union (Ece et al. 2018), Son Road, Nord 2000, Harmonoise, calculation of roadway traffic noise (CoRTN), ASJ RTN-model 2008, Federal highway administration (FHWA) model and CNOSSOS-EU (Garg and Maji, 2014). A comprehensive review of the aforementioned roadway traffic noise models by Garg and Maji, (2014) revealed that, considerations given to uncertainty computations in noise predictions as well as noise maps are limited, and suggested that, less arduous and time effective methods that also consider the uncertainty of the prediction of noise will be more appropriate for the stakeholders in providing a noise friendly environment. In addition to the non-generalization problem of the classical models due to the differences in local conditions such as road geometry, traffic volume and composition, the use of the classical models requires an in-depth understanding of the physical process and interaction between the traffic noise and the noise generators. The empirical models also proved to provide lower prediction accuracy than AI-based models in the prediction of nonlinear processes.

The limitations of the classical models give rise to the application of several AI-based models such as support vector machine (SVM), artificial neural network (ANN), random forest (RF), genetic algorithm (GA), decision trees (DT) and adaptive neuro fuzzy inference system (ANFIS) models for the prediction of roadway traffic noise, due to their accuracy and robustness in handling nonlinear processes like the traffic noise. For examples, Nedic et al., (2014) compared the performance of an ANN model with some statistical models for the estimation of highway traffic noise and the results affirmed the superiority of the ANN over other applied models. Kumar et al., (2014) employed ANN for modelling the roadway traffic noise of Punjab, India using the average speed, hourly traffic volume and heavy vehicle percentage as the model's inputs. A research conducted in Patiala, India evaluated and compared the performance of DT, RF, ANN and generalized linear model for roadway traffic noise estimation. The result showed that RF is more accurate and stable in the traffic noise prediction (Singh et al., 2016). Bravo-Moncayo et al., (2019) utilized three different machine learning approaches (ANN, SVM and multi-linear regression (MLR)) for the assessment of roadway traffic noise annoyance. The ANN model was found to be superior of the three models. Compared to the MLR and the SVM model, the modelling error in training phase was reduced by 42% and 35%, respectively and in testing stage the error was reduced by 24% for MLR and 19% for SVM model. Traffic noise in the hot climate of Sharjah, Dubai was modelled by ANN using five input variables namely mean speed, volume of heavy and light vehicles, road temperature and distance from the pavement edge. Comparing the efficiency of the developed ANN model with Ontario ministry of transport traffic noise model (ORNAMENT) and Basic Statistical Traffic Noise model (BSTN) in the prediction of roadway traffic noise proved superiority of the ANN model over the empirical models (Hamad et al. 2017). ANFIS model was also found to have higher prediction accuracy than FHWA, CRTN and RM models in a study performed by Sharma et al., (2018). ANN model has demonstrated superiority over two conventional roadway noise models (RLS90 and Criterion model) in the estimation of roadway noise in the mountainous city of Chongqing, China. The ANN had the least error of 1.60 dBA, while the RLS90 and Criterion had a forecasted error of 4.54 dBA and 6.70 dBA, respectively. The models input variables were traffic volume, speed, heavy-vehicle and road gradient (Chen et al. 2020). Ahmed and Pradhan (2019) developed an ANN model for both prediction and simulation of the propagation of roadway noise emission in a new expressway in Shah Alam, Malaysia. The model was found to have accuracy of 78.4% and an error of less than 4.02 dBA. Recently, Nourani et al., (2020a) developed a traffic noise model using an ensemble model that combines the outputs of AI-based models and a linear model where, the result of the ensemble approach provided higher accuracy than the single models. An emotional neural network (ENN) which is one of the recent generations of ANN that incorporates anxiety and confidence emotions into the ANN, was used by Nourani et al., (2020b) to model roadway traffic noise. The ENN led to higher accuracy in the prediction of roadway noise than the classic ANN and some common empirical noise models (CNR, RLS90 and BURGESS). Also, the study proved that dividing the traffic volume into sub-categories could enhance performance of the roadway traffic noise model up to 12% in the verification phase.

Although many AI-based models have already been utilized for roadway traffic noise prediction, and proved to be superior to both regression and empirical models, it is difficult to ascertain one particular model as a universal model that can predict roadway traffic noise with higher accuracy in all countries since different places have different traffic composition and characteristics. To overcome the constraints of the single models in modelling engineering processes like traffic noise, hybrid models have begun to attract attention of the researchers. The main aim of using hybrid models in prediction is to utilize each model's exclusive quality to capture different patterns in the data. Combining different models for predictions was found to be effective in improving the prediction accuracy in some other engineering and financial problems (e.g. see Nourani et al., 2011; Zhang et al., 2019). Besides, machine learning computing and modelling methods tend to be more demanding in under developing and developing regions where the economic stability and budget for environmental impact assessment is below the standard compare to developed countries. In contrast with this mentioned statement, the motivation of this research applies to both developed and developing countries as far as global warming and climate variables are concern.

In view of reported technical literature, to the best of the author's knowledge, there is no considered work in technical literature that employed application of the linear-nonlinear hybrid models for roadway traffic noise prediction. Based on the extensive bibliographic reported literature from Scopus database (1986-2021), there is need for substantial attention on road traffic noise modeling using the viability of AI based models. Figure 1a shows the main keywords of over 120 and likelihood occurrences, while Figure 1b indicating the important of this topic especially in African continent, Nigeria in particular. The conceptual approach of modeling road noise using new hybrid proposed in this work would be of interest and benchmarks to the researchers and scientist. This study presents the first application of Gaussian process regression (GPR) and a novel hybrid model for the prediction of roadway noise. The capability of the GPR model for prediction purposes has yielded reliable outcomes in several engineering problems (e.g. Bonakdari et al., 2019;Cai et al., 2020). The objective of this study was to propose novel linear-nonlinear hybrid models for estimation of roadway noise. This objective was achieved in two stages. First, development of five black box models (FFNN, SVR, GPR, BRT, MLR) and then development of four linear-nonlinear hybrid models (MLR-FFNN, MLR-SVR, MLR-GPR and MLR-BRT). To the best of the authors knowledge, there is no any

published study ay traffic
noise prediction

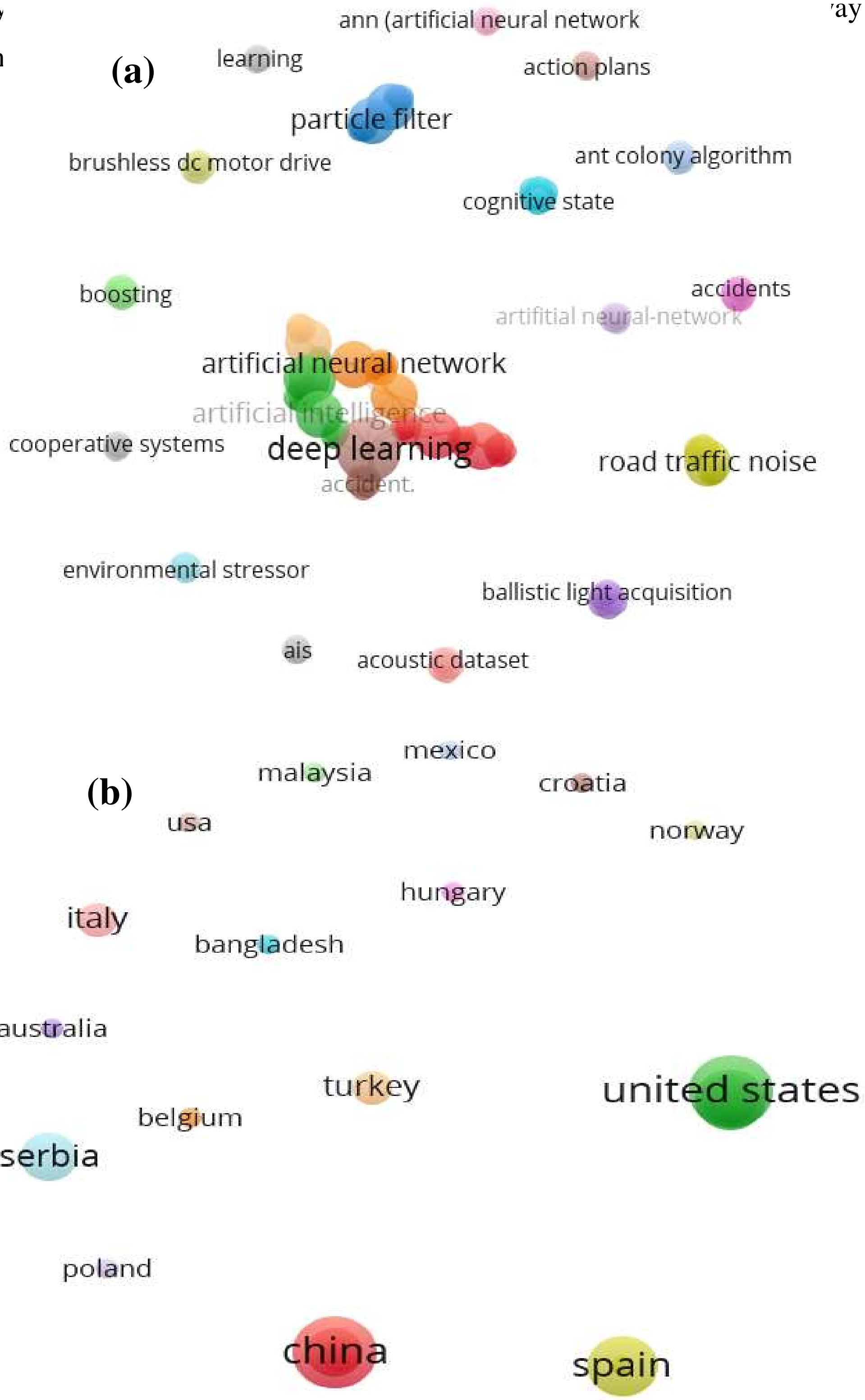


Figure 1: The algorithm results for the (a) Scopus database research for the surveyed keywords (b) the region/counties employed the road traffic noise research frequently

2. Material and Methods

2.1 Proposed methodology

The study was conducted in three stages (see Figure 2), selection of the relevant input variables was performed in the first stage. In the second stage, four different AI models (FFNN, GPR, BRT, SVR) and classical MLR model were used to model the roadway traffic noise at different sites in Nicosia, North Cyprus. Thirdly, the hybrid models were developed by summing up the predicted noise level from the MLR model and estimated residuals (MLR residuals) using the four AI-based models. The performances of the models were assessed and compared together.

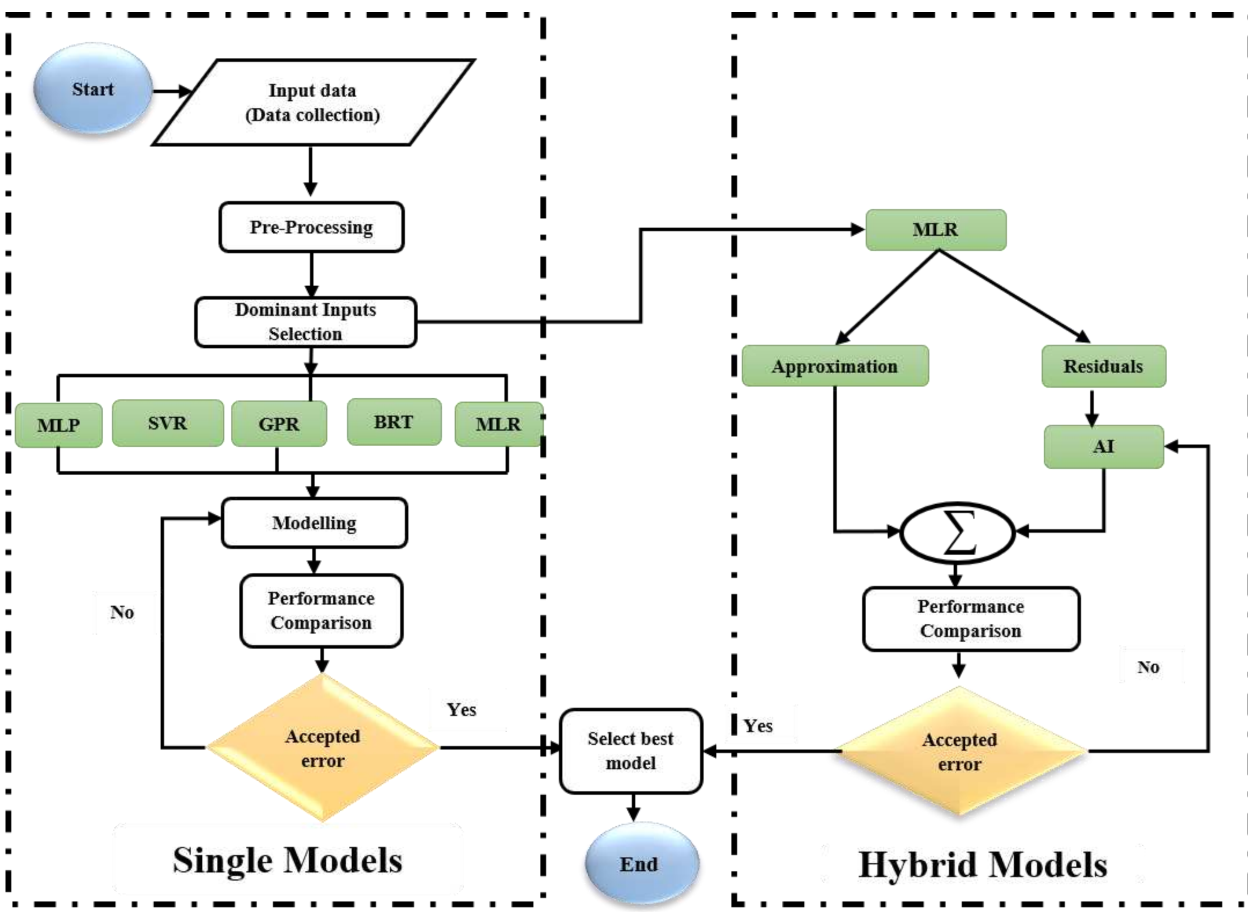


Figure 2: Proposed methodology for the prediction of roadway traffic noise

2.2 Data Description

For conducting the study, Sound Level Meter (SLM) with 0.1dBA resolution was used to record the roadway traffic noise level at some sites in Nicosia, North Cyprus (see Figure 3). The SLM was placed at 1.2m above the ground level at a distance of 3m from the road edge (see Figure 4). Simultaneously with the equivalent noise level (N), video camera was used to record the vehicular traffic. The traffic data consist of the vehicles' average speed (V), total traffic volume (Q) and traffic composition that is numbers of cars (C), medium vehicles (MV), bus (B), truck (HV) and motorcycle (MC). A total of 175 data samples from 12-observation sites on four road classes (local, collector, major arterial, and expressway) were recorded. The observations were taken in the morning (8:00-10:00) and evening (16:00-18:00) peak hours. Afternoon (12:00-14:00) observations were also taken at some strategic locations (sites 1-7 and 11). The relative humidity and wind speed at the time of data recording were less than 80% and 5 ms^{-1}, respectively.

The data collection sites were carefully sited such that, other noise sources were minimized to the lowest level. The observation sites were also located along straight tangent that are on relatively flat terrain and at a reasonable distance from any intersection to minimize the acceleration/deceleration effects. Table 1 summarizes the data recorded from the study sites. The noise level ranged from 56.3-80.5 dBA with 69.74 dBA as the mean value. The mean noise level at the 12 observation sites for both evening and the morning hours was greater than 55dBA that has been considered to be safe level in European countries (Ilgurel et al., 2016). The mean noise level for each of the road classes is shown in Figure 5, the expressway has the maximum average noise level followed by the collector road. The local road recorded the minimum noise level. This is attributed to the traffic volume and the average speed observed at the observation sites. A noise level of 80.5dBA was recorded in the evening hour which was recorded as the maximum noise level whereas the minimal noise level of 56.3dBA was observed during the morning hours. The highest traffic volume was recorded at site 5 during the evening hours while the highest vehicular speed was recorded along the express road (site 12) and minimum average speed and traffic volume were observed along the local roads (site 8) during the morning hours.

Table 1: Descriptive statistics of the recorded data

Parameter	Maximum	Minimum	Mean	Standard Deviation
Volume of cars (C)	981	36	405.43	227.12
Volume of medium vehicles (MV)	81	0	33.78	23.01
Volume of buses (B)	42	0	8.87	7.65
Volume of motorcycles (MC)	24	0	5.27	4.68
Volume of heavy vehicles (HV)	46	0	10.67	10.66
Number of honks (HK)	12	0	2.53	2.68
Velocity V (km/hr)	116	35	63.36	20.26
Noise level (dBA)	80.5	56.3	69.74	5.03

All observations made are for 15min intervals

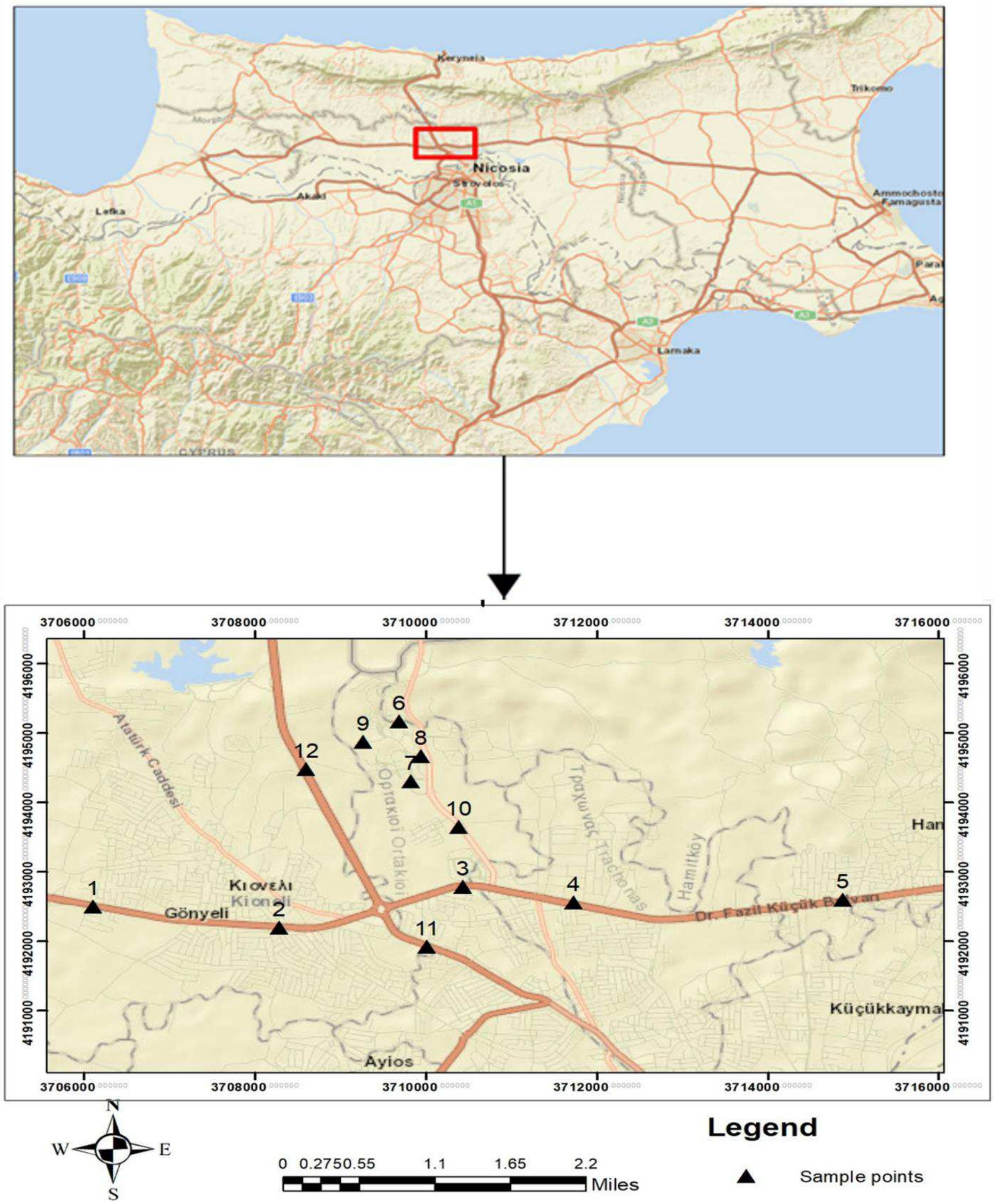
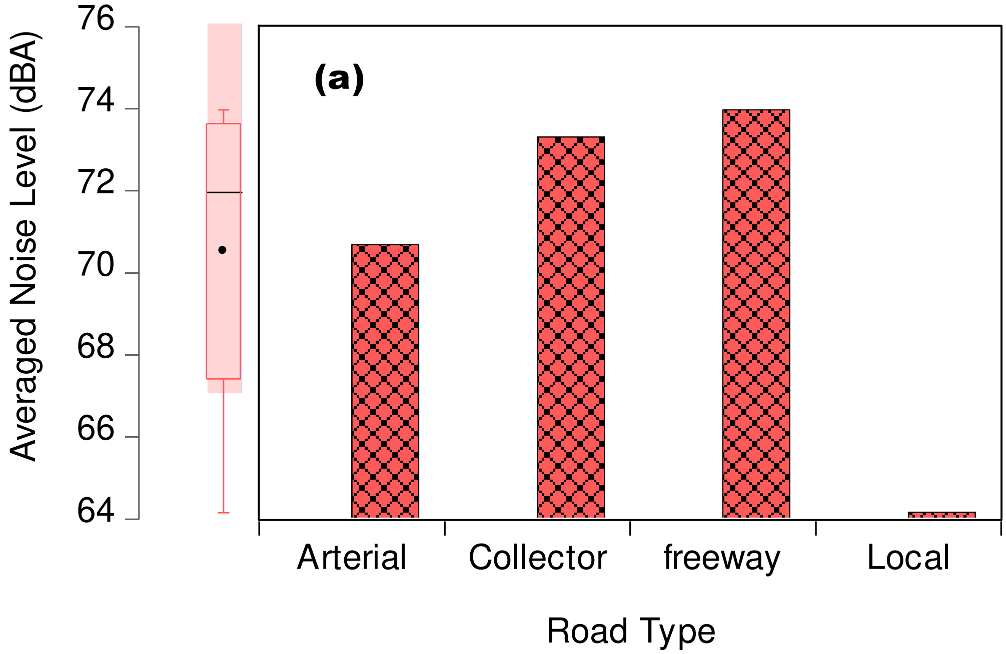


Figure 3: Study area showing sampling points

(a) **(b)**

Figure 4: Data collection (a) Camera setting at site 1 (b) Recording roadway traffic noise with SLM at site 5



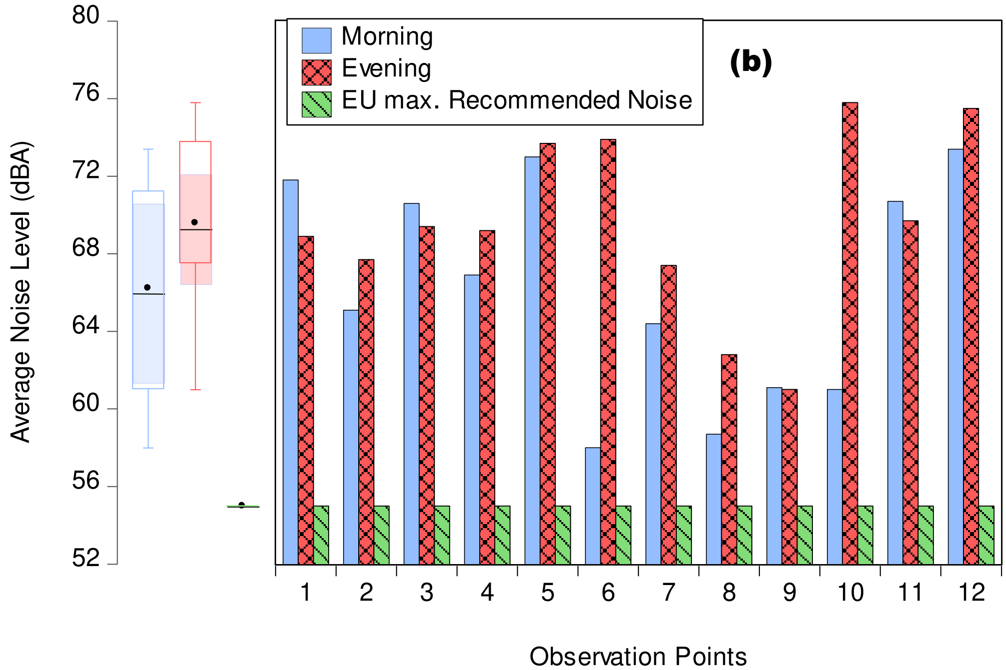


Figure 5: Average noise level (a) Road class (b) observation points

2.3 Dominant inputs selection

Mutual information (MI) method was used for the selection of the most appropriate and dominant input variables in the prediction of roadway noise. The MI measures the statistical nonlinear interactions between two variables (e.g. C and N), an MI value of 0 indicates that, there is no interaction between the two variables and a high MI value indicates strong nonlinear relationship (Nourani et al. 2015). Equation 1 was used for obtaining the MI value between the random parameters *x* and *y* (Yang et al., 2000).

$$MI\ (x,y)\ =\ H(x)\ +\ H(y)\ -\ H(x,y) \qquad (1)$$

whereas *H(x)* is the entropy function of *x* and *H(x, y)* is the joint entropy function of parameters x and y given as:

$$H(x,y)\ =\ -\sum m\epsilon x \sum n\epsilon y pxy\ (x,y)\ logpxy\ (x,y) \qquad (2)$$

P_{xy} *(x,y)* represents the joint probability distribution of *x* and *y*. The computed mutual information value between considered inputs and target are given in Table 2.

3. Artificial Intelligence Methods

Four AI-based models were employed for conducting the modelling and are briefly discussed in the below subsections.

3.1 Feed Forward Neural Network (FFNN)

FFNN is a type of supervised machine learning that may be trained to map the connection between input and output by altering the weights and biases between neuron elements. Despite ANN is available in different structures and training algorithms. The FFNN architecture consisted of three layers: an input layer, a hidden layer, and an output layer (see, Fig. 6). The hidden node processes the value given into the input layer and transmits the prediction to the output layer during the ANN process (Jahani and Mohammadi 2019). Furthermore, the training of the overall FFNN network was performed by adjusting the weight values to obtain the output with the lowest error. From the viewpoint of the optimization process, the FFNN training process is equivalent to the process to minimize a multivariable error function as a function of the network weights. The backpropagation algorithm adjusts the weight values by evaluating the gradient of the error function in relation to the weight values at each iteration (Kim and Singh 2014).

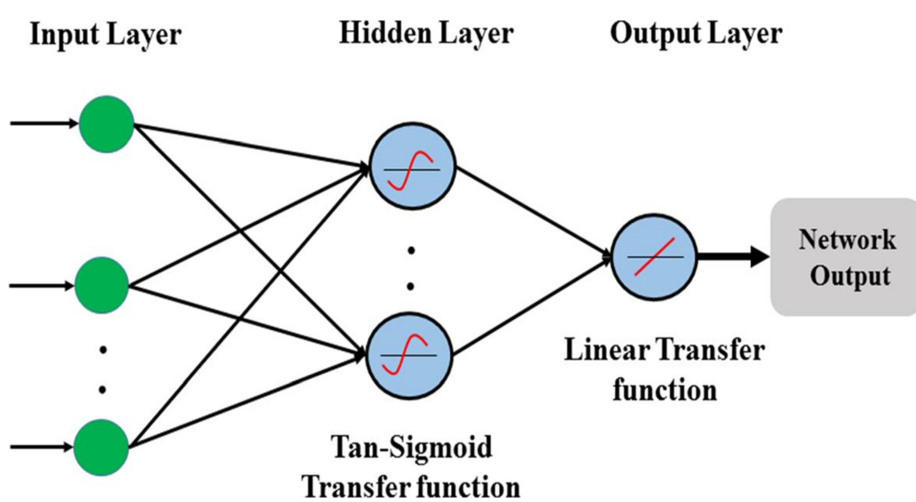


Figure 6: Structure of the three-layer FFNN Wang et al., (2015)

3.2 Support Vector Regression (SVR)

SVM was developed using statistical learning theory. The fundamental principle of the SVM execution in pattern recognition is the linear or non-linear mapping of the input vectors into a potentially higher dimension of feature space. The type of kernel function determines the mapping process. Then, an optimal hyperplane was built to achieve a maximum separation of two classes. In other words, SVM training was developed to address the issue of over-fitting and it excels at processing a large number of features (Vapnik, 1998). . For more details, the readers are referred to Wang et al., (2015) and Nourani et al., (2020b) about SVR modelling. Figure 7 gives the general structure of SVR model. The SVR equation can be expressed as (Wang et al., 2015):

$$f(x, \alpha_i, \alpha_i^*) = \sum_{i=1}^{N} (\alpha_i - \alpha_i^*) K(x, x_i) + b \quad (4)$$

Where x represents the input vector, α_i and α_i* are the Lagrange multipliers, $k(x_i, x_j)$ is the kernel function performing the non-linear mapping into feature space and b is bias term. Gaussian Radial Basis Function (RBF) kernel is the most commonly used kernel in the SVR and is given as:

$$k(x_1, x_2) = exp(-\gamma \|x_1 - x_2\|^2) \quad (5)$$

where, γ is the kernel parameter.

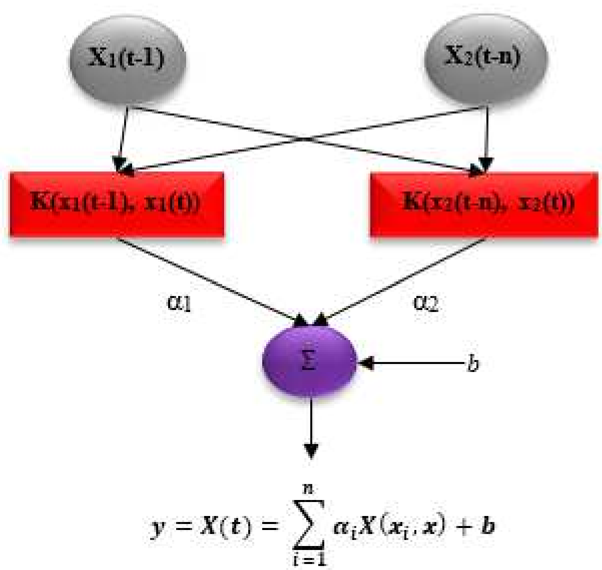


Figure 7: Conceptual structure of SVR model (Nourani et al. 2020a)

3.3 Gaussian Process Regression (GPR)

Gaussian Process Regression (GPR) is a non-parametric technique that is used to model random complex systems. The flexibility of the GPR method in providing uncertainty representation makes it more desirable in the prediction of many engineering problems (Rasmussen 2004). The GP is a stochastic process of which finite sub-collection of random variables has a multivariate Gaussian distribution (Cai et al. 2020). The general expression of the GPR model relating the explanatory vector (*x*) and the response (*y*) is given by:

$$y_i = f(x_i) + \varepsilon \tag{6}$$

In Equation 6, $f(x_i)$ stands for an arbitrary function that maps the inputs into the corresponding outputs, ε represents the regression error having an identically distributed Gaussian function with mean and variance values of zero and σ^2, respectively.

The function *f(x)* for any unobserved pair *(x*, f*)* in which *f* is the response and *x* is the explanatory parameter is obtained by:

$$\begin{bmatrix} f \\ f^* \end{bmatrix} \sim \mathrm{N}_{n+1}\left(0, \begin{bmatrix} K(X,X) & k(X,x^*) \\ k(x^*,X) & k(x^*,x^*) \end{bmatrix}\right) \tag{7}$$

In Equation 7, *K (X, X)* represents the matrix of covariances (*n* x *n*) for all samples in the calibration data, *k* (X, x*) stands for vector of covariances (*n* × 1) between the point *x** and calibration data. *k(x*, x*)* is the variance at point *x**. In the classic regression, the mean (*f*) is derived from *f* then integrates to *f**:

$$P(f * | x *, X, f) = N|(k(x *, X)K(X,X)^{-1} f, k(x *, x *) - k(x *, X)K(X,X)^{-1} k(X, x *))$$

(8)

Equation (8) expresses X and *f* by maximizing the joint probability of *f** conditional on *x** to obtain the *f**.

When using data that is noisy, it should be supplemented by a model for the observation error. Hence, Equation (7) is converted into:

$$\begin{bmatrix} f \\ f^{*} \end{bmatrix} \sim \mathrm{N}_{n+1}\left(0, \begin{bmatrix} K(X,X) + \sigma^{2}\mathrm{I} & k(X,x^{*}) \\ k(x^{*},X) & k(x^{*},x^{*}) \end{bmatrix}\right) \qquad (9)$$

consequently, the conditional likelihood and the variance change to

$$f(x^{*},) = k\,(x^{*}, X)\,(K\,(X,X) + \sigma^{2}\, I)^{-1} f \qquad (10)$$

and

$$Cov\,(f(x^{*})) = k\,(x^{*}, x^{*}) - k\,(x^{*},X)\,(K\,(X,X) + \sigma^{2}\, I) - k\,(x, x^{*}) \qquad (11)$$

where I stands for identity matrix and σ^2 represents variance of the measured error (Bonakdari et al. 2019).

3.4 Boosted Regression Tree (BRT)

The BRT is a unique method for prediction and classification combining both a machine learning approach and a statistical technique. The BRT combines several models and fit them into single model for improving performance of the single models in prediction problems (Youssef et al. 2016). The method does not require any data transformation before fitting the complex nonlinear pattern of the dataset and establishing the interaction between the target and input variables (Elith et al. 2008). This advantage of the BRT makes it suitable for modelling natural processes with complex nonlinear relationships. Information in decision trees is represented in distinctive way that is easy to visualize which gives it several advantages. In the BRT, missing data in the predictor variables are modified using surrogates (Elith et al. 2008). Another advantage of all decision trees including the BRT is their insensitivity to outliers. Boosting and regression are the two algorithms used in the BRT models. Boosting is a technique used for enhancing prediction accuracy of a model based on the idea that, it is easier to find many rough rules of thumb than to find a single and highly accurate prediction rule (Youssef et al. 2016). Fitting multiple regression trees in the BRT overcomes the deficiency of the single regression trees in predictions. The Regression Learner of Matlab (2019b) was employed for developing the BRT model in this study. For a typical predictive learning system consisting of a set of predictors of different variables X= $\{x_1\ldots, x_n\}$ and a response variable y, a BRT for function approximation could be applied. For example, using a training sample $\{y_i, X_i\}$, i =1…, N of known y and X values. The aim is to determine the function F^*(X) (Equation 12) that fits X to y, such that the anticipated value of the identified loss function is minimized over the joint distribution of all values of *X* and *y*. In gradient boosting regression, the function is approximated using Equation 13.

$$F^*(X) = \psi(y, F(X)) \quad (12)$$

$$F(X) = \sum_{m=0}^{M} F_m(X) = \sum_{m=0}^{M} \beta_m g(X; \alpha_m) \quad (13)$$

Where g (X; α_m) stands for the regression tree at a specific node, β_m are the expansion coefficients, αm explains the tree parameters, m=1…, M. The X space is divided into N-disjointed regions {R_{nm}} for each iteration m, n=1…, N and distinct constant are estimated in each iteration (Suleiman et al. 2016). The following steps are employed for implementing the BRT algorithm:

1. Initialize *F*(X) to be a constant
2. Do the following steps for values of m from 1 to M:
   a. Compute the residual error $r = -\left[\partial\psi y_i, \frac{F(X_i)}{\partial F(X_i)}\right] F_m(X) = F_{m-1}(X), i = 1, \ldots, N$
   b. Without replacement, select randomly *p* ×N samples from the calibration data.
   c. To obtain the approximate α_m value of $\beta g(X; \alpha)$, fit the r values computed is step 2a into a least squares regression trees with K terminal nodes using the randomly selected observations in 2b.
   d. Minimize the loss function $\psi(y, Fm - 1(X)) + \beta g(X; \alpha_m)$ to obtain the approximate values of β_m.
   e. Update $F_m(X) = F_{m-1}(X) + \beta_{mg}(X; \alpha_m)$
3. Calculate $F_m(X) = \sum_{m=0}^{M} F_m(X)$

For the avoidance of overfitting problems expected in the BRT models, a learning rate λ parameter that controls the contribution of each regression tree is added to keep the condition under control by moderating the calibration process of the regression trees as shown in Equation 14. There is a strong interaction between λ and number of iterations M. For convergence of the calibration error, more iterations are required for smaller values of *m*. Setting the λ to a small constant value and choosing fewer number of iterations has been recommended by Hastie et al., (2011) for obtaining better test error.

$$F_m(X) = F_{m-1}(X) + \lambda \beta_{mg}(X; \alpha_m) \quad (14)$$

3.5 Multi Linear Regression (MLR)

The most commonly used method for the prediction and analysis of engineering problems is the MLR. It helps understand the linear dependency between the predictor and the dependent variables. It explores the interaction between the variables and describes the relationship between them by keeping the independent variables fixed and varying one (Doğan and Akgüngör 2013). The n regressor variables and the dependent variable y can be correlated by Nourani et al., (2020a):

$$y = b_0 + b_1 x_1 + b_2 x_2 + b_3 x_3 + \cdots + b_i x_i + \xi \tag{15}$$

In Equation 15, x_i represents value of the i^{th} predictor, $b_i$stands for coefficient of the i^{th} predictor, b_0 is the constant of regression and ξ is the error term.

3.6 Proposed Hybrid Methodology

In many real-life problems such as the roadway traffic noise prediction, a linear or a nonlinear interaction may exist between the predictor variables and the roadway traffic noise level. As a result of this complex nature, the application of linear models (such as MLR, ARIMA) for such process may not be adequate. On the other hand, nonlinear models (such as ANN, SVR etc.) despite their advantage in modelling complex problems are not appropriate for all circumstances and may yield errors especially in modelling data with a linear pattern. Therefore, it is not appropriate to blindly apply nonlinear models to any data without pre-processing of the data. For example, a spatial data pre-processing (e.g. spatial clustering) should be employed for modelling processes that shows trend in space before developing main models. Likewise, temporal data pre-processing may improve the model efficiency for processes which include seasonal and non-stationary characteristics. Therefore, since it is difficult to know the characteristics of the data in a real problem completely, hybrid modeling can be a good methodology for practical use and by combining different models; different aspects of the underlying patterns may be captured (Nourani et al., 2011). In this study, a hybrid model was developed by combining the predicted values from a linear model (MLR) and estimated residuals (error) by a nonlinear model (AI-based). The proposed hybrid model can be expressed as:

$$y_i = L_i + N_i, \tag{16}$$

where y_i is the observed noise level; L_i and N_i are the linear and nonlinear parts of the traffic noise, respectively. The development of the proposed linear-nonlinear hybrid model involved three steps (Figure 8). For the first step, a linear model is created via MLR and the residuals are computed using:

$$r_i = N_{obs(i)} - N_{pre(i)}. \tag{17}$$

where the residual r_i is estimated by the MLR models, $N_{obs(i)}$ and $N_{pre(i)}$ present the observed noise level, and predicted noise level by MLR model, respectively. In the second stage, the residual (r_i) which contains only the nonlinear part of the traffic noise that was not captured by the MLR, is passed through a nonlinear kernel of AI model, (e.g., FFNN, SVR, BRT and GPR) for capturing the nonlinearity of the data. Lastly in step three, the result obtained from the nonlinear model is combined (summed up) with the output of the MLR model obtained in step 1 to give the predicted noise level by the hybrid model. The final traffic noise computed by the hybrid model is given by Equation 18. By combining the MLR and AI-based models in roadway traffic noise prediction, the MLR will effectively capture the linear pattern in the data and the AI-based models will capture the nonlinearity of the data there by coming up with a model that has higher prediction accuracy than both MLR and the AI-based models as hinted by Nourani et al., (2011).

$$Y_{pre} = N_{pre} + r_{pre} \tag{18}$$

Where Y_{pre} is the predicted noise level, N_{pre} stands for the approximated noise level obtained by MLR and r_{pre} is the predicted residual obtained using the AI model.

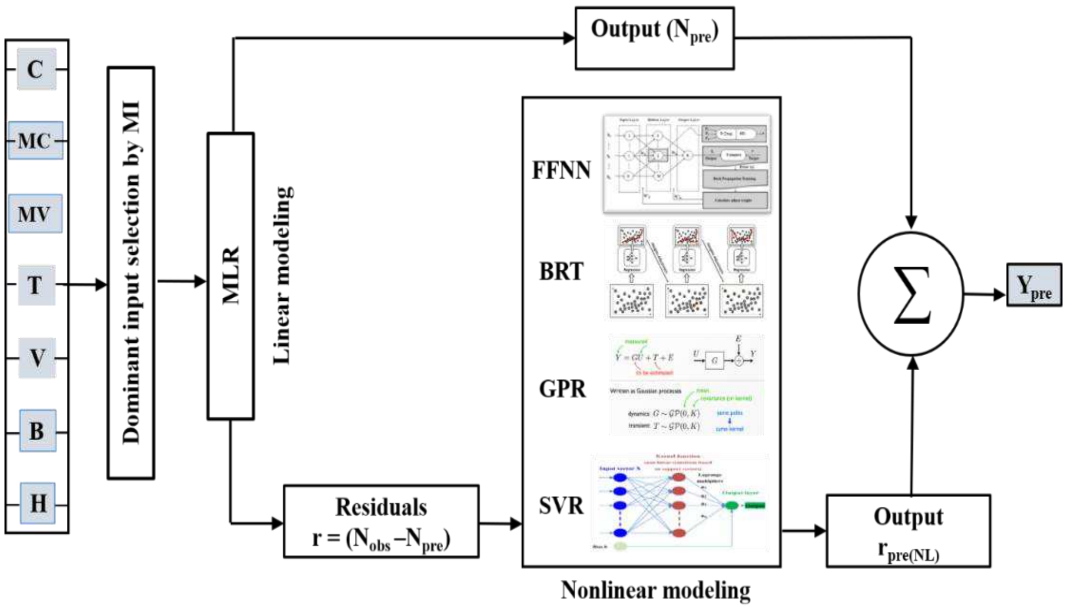


Figure 8: Proposed linear-nonlinear hybrid model

3.7 Model Validation

The main purpose of using data-driven models for prediction problems is to achieve a reliable result that is difficult to obtain using the classical methods without prior knowledge and deep understanding of the concept. But, due to overfitting problems in many data driven models, the model's performance at the calibration stage is not always coherent with its performance at the verification stage, which makes it impossible to obtain accurate prediction results for other unseen dataset. This makes it necessary to validate the models for overcoming the overfitting issues. Despite the fact that, hybrid models handle the overfitting problems much better than the traditional feed forward neural network, because the main part of the model is MLR (linear model) which is not so sensitive to overfitting issue, it may also experience overfitting issues as a result of fewer observation samples for training the model. Different types of validation process exist in the literature (cross validation, holdout validation and leave one out validation etc.) but the k-fold cross validation was employed in this study. In this type of validation mechanism, the dataset is portioned into equal k-number of subsets. The calibration of the model is done using k-1 subsets and remaining subset is used for the verification. The procedure is repeated for k times until all the k-subsets are used for the calibration and verification in alteration. The final performance is obtained by computing the average value of k- subsets performances in verification stage. One of the key benefit of using the k-fold cross validation is that the calibration and the verification subsets are independent (Sharma et al. 2018). Efficiency in the data usage could also be achieved through the cross validation. Considering the 4-fold cross-validation, the data set (normalized) is divided into two (calibration=75 % and validation=25 for developing the models. The data size determines the k values to be used usually ranging from 2-10.

3.8 Data pre-processing and performance evaluation

In order to obtain a better result, prior to the development of the data driven models, data preparation such as normalization, standardization etc., are required. In this study, normalization was employed and all data (inputs and the target parameters) were normalized between the values of 0 and 1. The normalization was done to bring all the input and the target parameters to the same range and to also remove the dimensions of the data. This helps prevent that data with higher numeric values to dominate over those with lower values (Nourani et al., 2019a). Normalization also improves the model's accuracy by reducing the complexity, computational requirement, redundancy in the data and also time required to attain the global minima. The data were normalized using:

$$N_{norm} = \frac{N - N_{min}}{N_{max} - N_{min}} \tag{19}$$

Where N, N_{max} and N_{min} represent the values of observed, maximum and minimum roadway noise levels, respectively, while N_{norm} indicates the normalized roadway noise

The efficiency of the developed models in predicting the equivalent noise level was evaluated using four different evaluation criteria namely the mean absolute error (MAE), root mean square error (RMSE), Nash-Sutcliffe efficiency (NSE) and the relative root mean square error (rRMSE). The NSE values ranges from -∞ to 1 and it is a parameter that indicates how well the model fits the observed noise level. A perfect model has an NSE value of 1 and the model efficiency decreases as the value moves far from 1 and vice versa Nourani et al., (2020a). RMSE as one of the best measures for computing the model's performance was used for measuring the average error produced by the models. The RMSE value ranged between 0 and +∞ and is zero in the best model (Nourani and Sayyah 2012). The MAE construes the goodness-of-fit of the model regardless of the sign of the prediction error between observed and predicted noise level values just like RMSE. MAE was used in the study for evaluating the deviations of the predicted noise level from the observed values in an equal way regardless of the sign since the RMSE is suitable for estimating errors with a normal distribution which may not be satisfied by all proposed models (Bonakdari et al. 2019). Finally, rRMSE was also used, which could be evaluated based on the defined ranges: Excellent for rRMSE values less than 10%, Good for values between 10% and 20%, Fair for rRMSE values between 20% and 30%, and Poor if rRMSE value is greater than 30% (Rabehi et al. 2020). The performance evaluations mentioned are computed using Equations 20 -23, respectively.

$$NSE = 1 - \frac{\sum_{i=1}^{n}\left(N_{obs_i} - N_{pre_i}\right)^2}{\sum_{i=1}^{n}\left(N_{obs_i} - \overline{N_{obs_i}}\right)^2} \tag{20}$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}\left(N_{obs_i} - N_{pre_i}\right)^2}{n}} \tag{21}$$

$$\mathrm{MAE} = \frac{\sum_{i=1}^{N}\left|N_{obs_i} - N_{pre_i}\right|}{N} \tag{22}$$

$$\mathrm{rRMSE} = \frac{\sqrt{\frac{\sum_{i=1}^{n}\left(N_{obs_i} - N_{pre_i}\right)^2}{n}}}{\frac{1}{n}\sum_{i=1}^{n}\left(N_{obs_i}\right)} \times 100 \tag{23}$$

where, n is the number of observations, $\overline{N}_{obs}$ is the mean observed noise level, N_{obs} is the observed noise level, and N_{pre} is the predicted noise level.

4. Results and Discussion

4.1 Results of single models

For all data driven models, selection of the dominant input variables is very important for obtaining a satisfactory result as fewer inputs might not accurately model the process and too much inputs might increase the complexity of the model. In this study, MI between the inputs and target was computed and used for dominant variables selection purpose and in this way, C, V, MV, HV and B were found to be the most dominant parameters among the potential inputs with MI values greater than 1 as shown in Table 2. The result of the MI value showed that C is the most important parameter for the prediction of roadway noise in Nicosia followed by V, and are ranked 1^{st} and 2^{nd}, respectively. This result was supported by findings of Agarwal and Swami (2011) where volume of light vehicles were suggested to be the most important parameters in modelling roadway noise. Likewise, Covaciu et al. (2015) and Vijay et al. (2015) stressed on the importance of speed for the generation of roadway noise. MV and HV which are ranked 4^{th} and 5^{th}, were reported to have significant influence on roadway noise generation by Gökdag (2012) and Vijay et al. (2015), respectively. The MC and HK got a low MI values (<1) and were found to be statistically insignificant according to the student t-test. Despite the fact that other studies (e.g. Vijay et al. 2015) showed increase in roadway noise as a result of honking. In this study, number of honks was found to be insignificant due to fewer number of honks during the observation period. The low number of the honk is because the data collection was done along a straight road section with less conflict. Also, volume of MC is low due to the nature of the traffic composition in Nicosia. Local parameters such as road geometry and traffic composition of a region or country has significant influence of the vehicular class and other parameters that maybe dominant for the generation of roadway noise in that region. The modelling part of the study was done in two stages after determining the dominant input variables.

Table 2: Mutual information (MI) value between the potential inputs and target (traffic noise)

C	**V**	**MV**	**HV**	**B**	**MC**	**HK**
1.5686	1.4785	1.3538	1.2001	1.0608	0.8229	0.5962

In the first stage, five data driven models including four AI-based models (BRT, FFNN, GPR, SVR) and a classical model (MLR) were developed for the prediction of roadway traffic noise, individually. The performances of these models were evaluated using four performance criteria (NSE, RMSE, MAE, CC) and the results are presented in Table 3. It should be noted that, several models were developed with each of the AI-technique using different structure, training algorithms and kernel functions but only the best models are reported in the Tables. For the FFNN model, the best result was obtained using 5-8-1 structure (8- neurons in the hidden layer) trained with the Levenberg Marquardt algorithm and tan-sigmoid activation function. The best models for the SVR, BRT and GPR were obtained using RBF kernel, least square boost algorithm and squared exponential kernel which is the most commonly used kernel for GPR models (Athavale et al. 2019), respectively. All the AI-based models led to reliable performance in the prediction of roadway traffic noise than the linear model (MLR). This is because, the AI based models have higher ability in modeling convoluted processes such as roadway traffic noise in addition to better generalization ability than the linear models (Nourani et al. 2019b). It can be seen that, the BRT model outperformed all other models used in this study in the prediction of the road noise by providing higher NSE value and lowest error (RMSE, MAE) at verification stage.

Comparing the performances of the AI-based models with the MLR model indicated that, the BRT model has an improved prediction accuracy over the MLR model by up to 20.9% in the verification step, GPR up to 16.9%, FFNN up to 12.6% and lastly the SVR model which has an improved performance over the linear model by 10.3%. The BRT demonstrated higher prediction capability than GPR, FFNN and SVR models at both calibration and verification stages with NSE, RMSE and MAE values of 0.8679, 0.0852, 0.0626 respectively, at the verification stage. The ability of the BRT to model with high accuracy comes by ensemble of different regression tress and its ability to fit complex nonlinear relationships and automatically addressing the interaction effects between the predictions. The results shown in Figure 9 show that the BRT fits the data better than all other models in verification stage. The data are more compacted along the diagonal line better than the other models which indicates better goodness of fit than other data driven models. Followed by the BRT is the GPR model, even though this is the first study to apply GPR model in vehicular traffic noise prediction, a similar outcome where GPR outperformed ANN, SVR, and POD was also reported by Athavale et al., (2019) in a study to compare the prediction capability of different data driven models for temperature time series prediction. The GPR gets its high prediction ability from its flexibility to provide uncertainty representation (Cai et al. 2020). In terms of the model's accuracy, the FFNN model was more accurate than all other data driven models with least percentage increase in the NSE values between the calibration and verification stages, followed by the GPR model. Contrary to the findings by the Fan et al., (2018) where SVR was found to be more stable than the tree-based ensemble algorithms in the prediction of daily evapotranspiration, the SVR was found to be least stable for prediction of roadway traffic noise.

The efficiency of the different models (with regards to NSE values) in both calibration and validation stages were compared by radar charts as shown in Figure 10. In addition to the radar chart, Taylor diagram was also used to compare the models' performances (see Figure 11). The Taylor diagram compares different statistical performance metrics (RMSE, correlation and the standard deviation) of the models. In the Taylor diagram, the azimuthal position gives the correlation between the actual and the computed values. The RMSE values are directly proportional to the distance between the observed and the predicted fields having same unit with the standard deviation. For any increase in correlation, the value of the RMSE is decreased. The standard deviation of the pattern increases with increasing radial distance measured from the origin (Taylor 2001). A model is said to be a perfect model by reference point when its correlation coefficient is 1 (Yaseen et al. 2018). If the standard deviation of the computed values is greater than the standard deviation of the measured values, then it may lead to overestimation and vice versa. However, considering the rRMSE values of the models at the verification stage (>20%), it shows that all models have fair performance with the exception of the BRT model which led to almost good performance. It is clear that, there is a need to improve the modelling performance of the process. To this end, the following hybrid models were for prediction of the roadway traffic noise.

Table 3: Performance of single models for prediction of roadway traffic noise level

Model	Calibration				Verification			
	NSE	RMSE*	MAE*	rRMSE*	NSE	RMSE*	MAE*	rRMSE*
FFNN	0.7857	0.0754	0.0534	13.6005	0.7850	0.1325	0.1035	23.9084
SVR	0.8406	0.0650	0.0417	11.7299	0.7619	0.1394	0.0952	25.1564
BRT	0.9110	0.0592	0.0464	10.6796	0.8679	0.0852	0.0626	15.3848
GPR	0.8687	0.0590	0.0389	10.6452	0.8282	0.1184	0.0882	21.3712
MLR	0.6707	0.0934	0.0724	16.8603	0.6586	0.1669	0.1214	30.1236

*No unit for normalized data

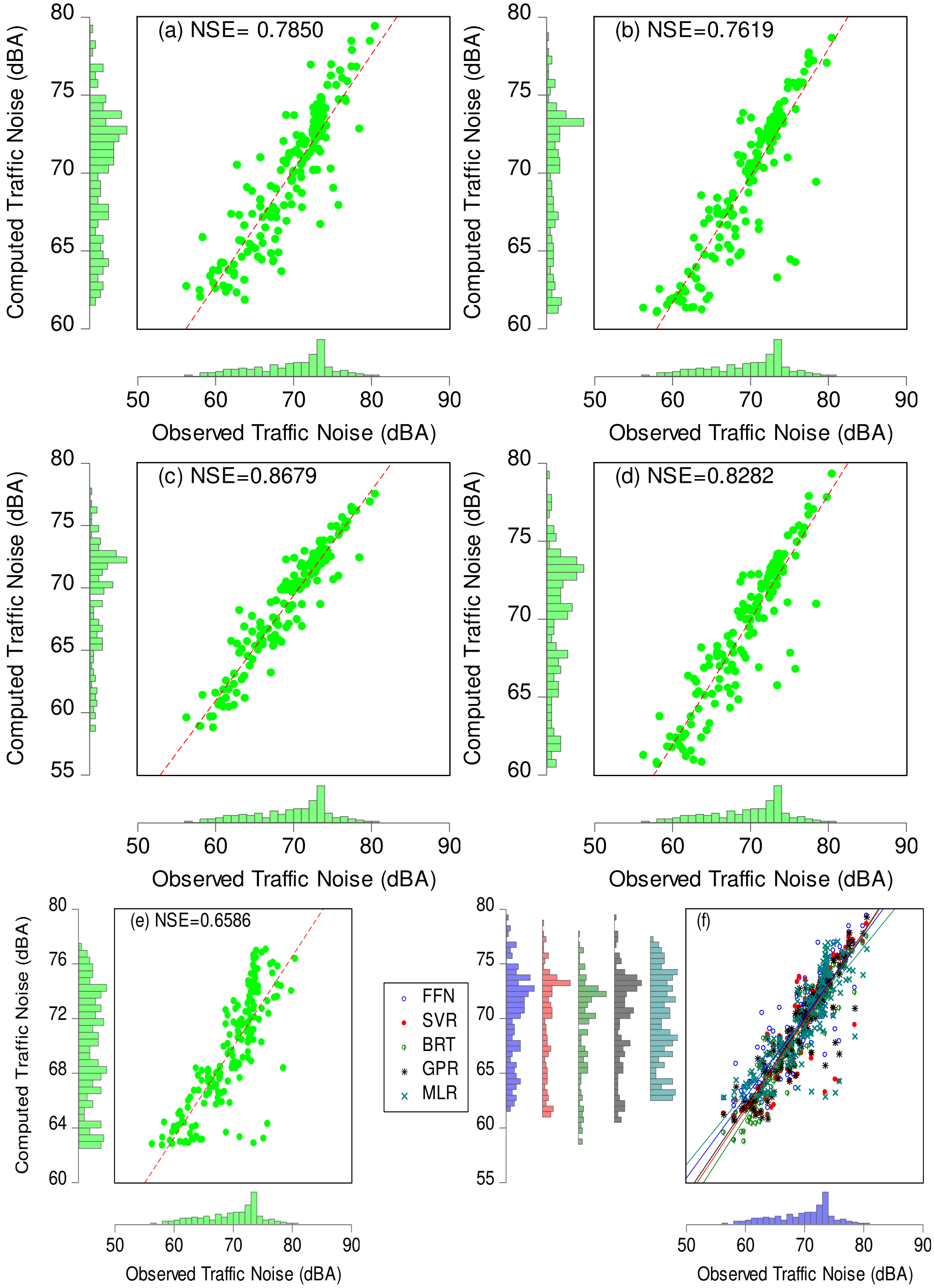


Figure 9: Scatter plots of observed and computed roadway traffic noise levels in verification stage obtained by a) FFNN, b) SVR, c) BRT d) GPR, e) MLR, and f) overall models

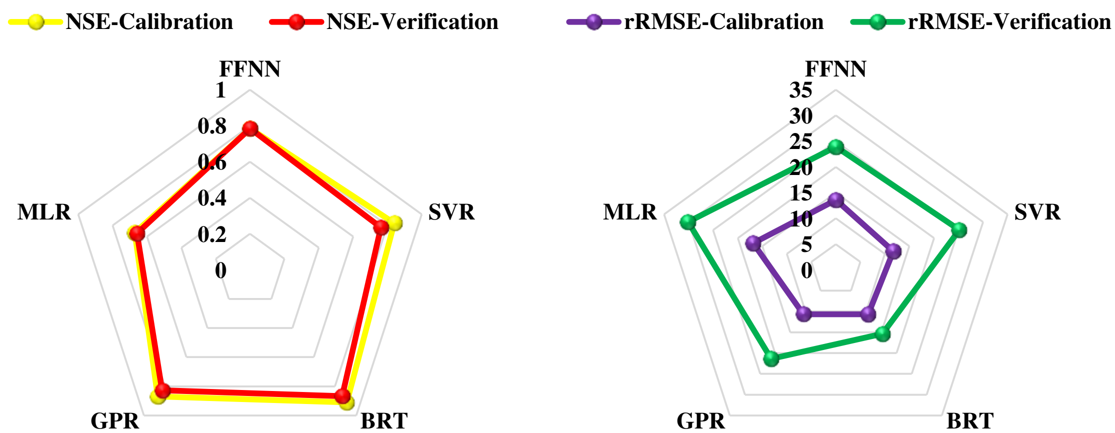


Figure 10: Radar Plot showing NSE and rRMSE values for different models in calibration and verification stages

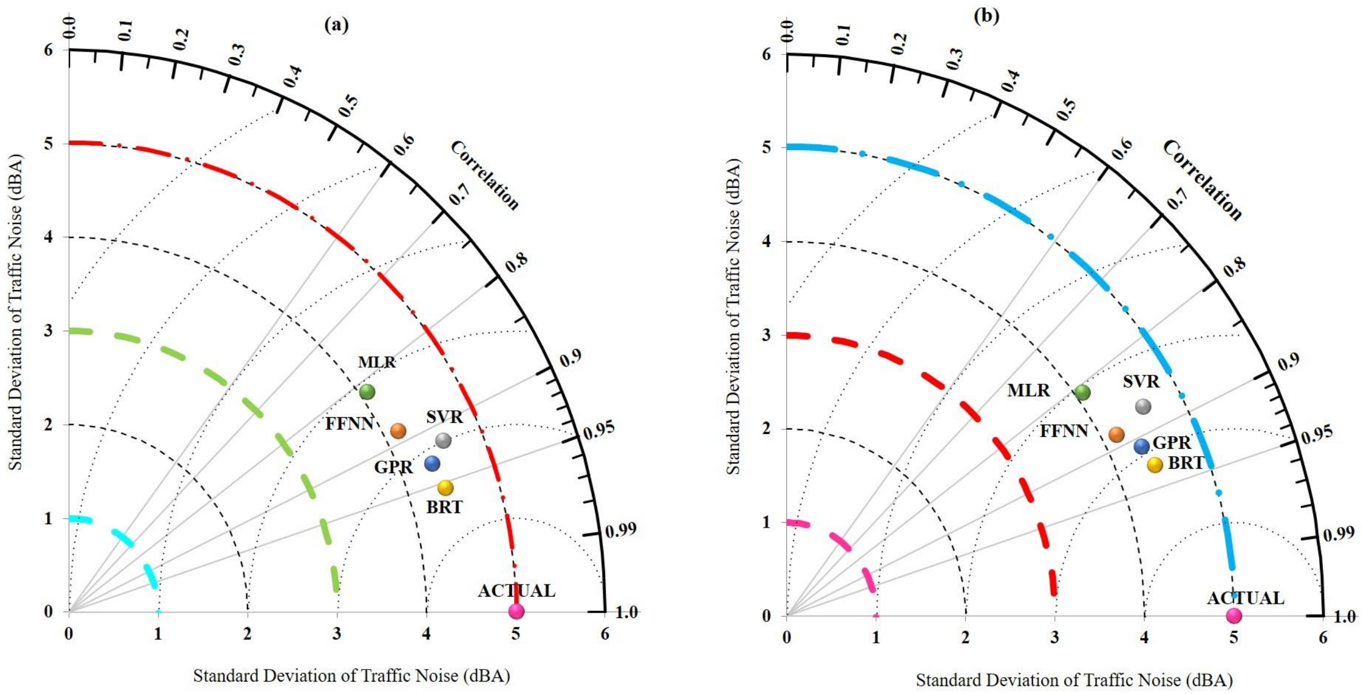


Figure 11: Taylor diagram representing different statistical parameters of the models in steps of (a) Calibration (b) Verification

4.2 Results of hybrid models

For enhancing the prediction ability of the single models in this study, four different linear-nonlinear hybrid models were developed in the second part of the study where the results of the hybrid models are shown in Table 4. The results of the hybrid models demonstrated increased performance in the prediction of roadway traffic noise with regard to the single models. Similar results were obtained by Nourani et al., (2011) where SARIMAX (Seasonal Auto Regressive Integrated Moving Average with exogenous input)-ANN model outperformed both SARIMAX and ANN models in daily and monthly rainfall-runoff modelling at both calibration and verification stages. Zhang et al., (2019) also found that linear-nonlinear hybrid (autoregressive integrated moving average (ARIMA)-SVR) could model emergency patient flow with higher accuracy in terms of MAPE, MAE, and RMSE than both the ARIMA (linear) and the SVR (nonlinear) models. The MLR-GPR model demonstrated higher performance at the verification stage than all linear-nonlinear hybrid models with NSE value of 0.9312, followed by MLR-BRT (0.9100), then MLR-FFNN (0.8845) and finally MLR-SVR (0.8723). Performance evaluation of the linear-nonlinear hybrid models using the rRMSE showed that, all the hybrid models have excellent to good performance with MLR-GPR having the highest accuracy with rRMSE value of 7.4% (excellent). Table 5 clearly indicates that the hybrid modelling improved the performance of the nonlinear models by up to 10.30% for the AI models and up to 27.26% for linear models. The predictive performance of the hybrid models is presented graphically using the Taylor diagram (see Figure 12) and Radar plots (Figure 13). It can be seen clearly that, the MLR-GPR hybrid model was the best model outperforming all single and linear-nonlinear hybrid models. Comparing the absolute error of the hybrid models in the roadway traffic noise prediction using the box plot in Figure 14 revealed higher accuracy of the MLR-GPR hybrid model. The MLR-GPR model has the least forecasted mean absolute error (0.85 dBA) than all the hybrid models, making it reliable for the estimation of roadway traffic noise. As a result, the hybrid model could be used for enhancing the performance of the non-linear models. The results from the hybrid models could be integrated for development of a more accurate and reliable traffic noise maps that will in turn help the stakeholders in providing a sustainable mitigation measure for reducing the peoples' incessant exposure to the traffic noise. The use of pavement materials with suitable textures during the construction, car sharing and the use of electric cars are some of the sustainable tools in providing a noise healthy environment. The aforementioned statement could be justified by considering the comparative analysis of hybrid scatter plot presented in Figure 15 and that of single model above.

Table 4: Performance of hybrid models for prediction of traffic noise level

Models	Calibration				Verification			
	NSE	RMSE*	MAE*	rRMSE*	NSE	RMSE*	MAE*	rRMSE*
MLR-FFNN	0.9657	0.0529	0.0450	9.9861	0.8845	0.0553	0.0422	9.5447
MLR-SVR	0.9610	0.0564	0.0465	10.1826	0.8723	0.0582	0.0470	10.5005
MLR-BRT	0.9440	0.0676	0.0398	8.8154	0.9100	0.0488	0.0529	12.1971
MLR-GPR	0.9793	0.0411	0.0350	7.7069	0.9312	0.0427	0.0347	7.4249

*No unit for normalized data

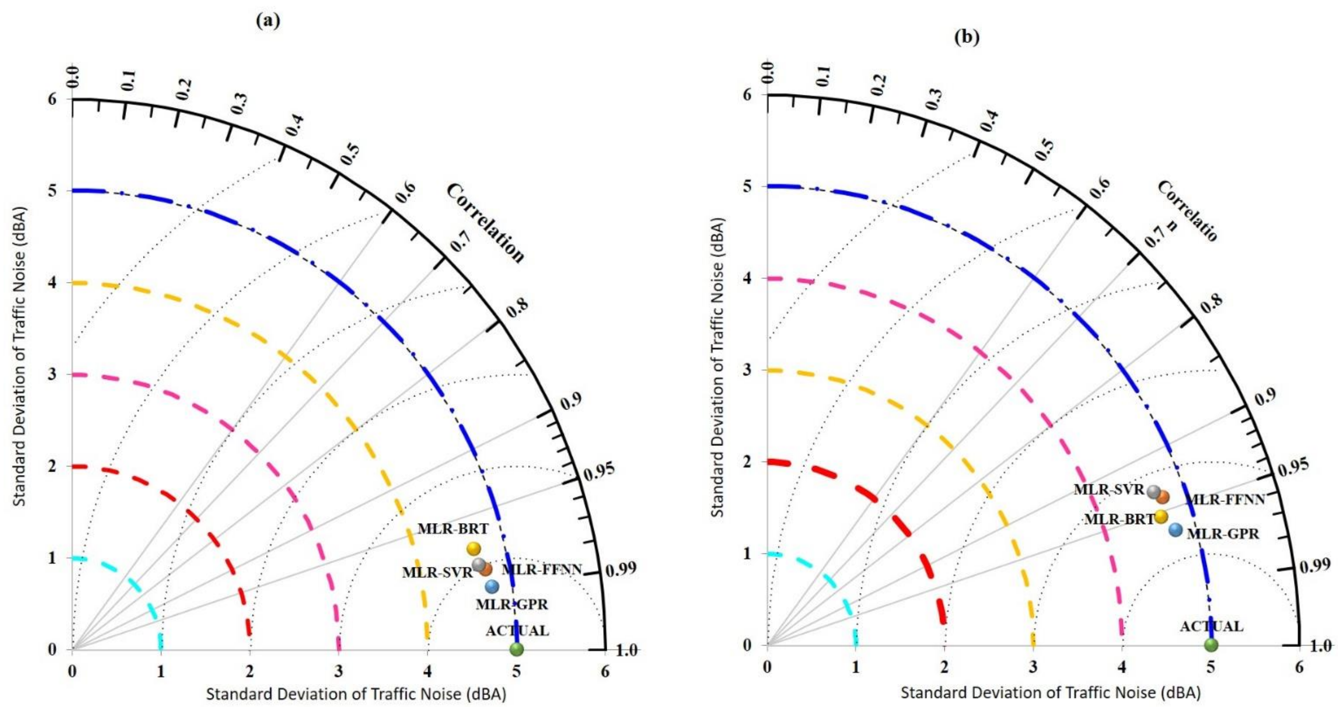


Figure 12: Taylor diagram representing different statistical parameters of the hybrid models in steps (a) Calibration (b) Verification

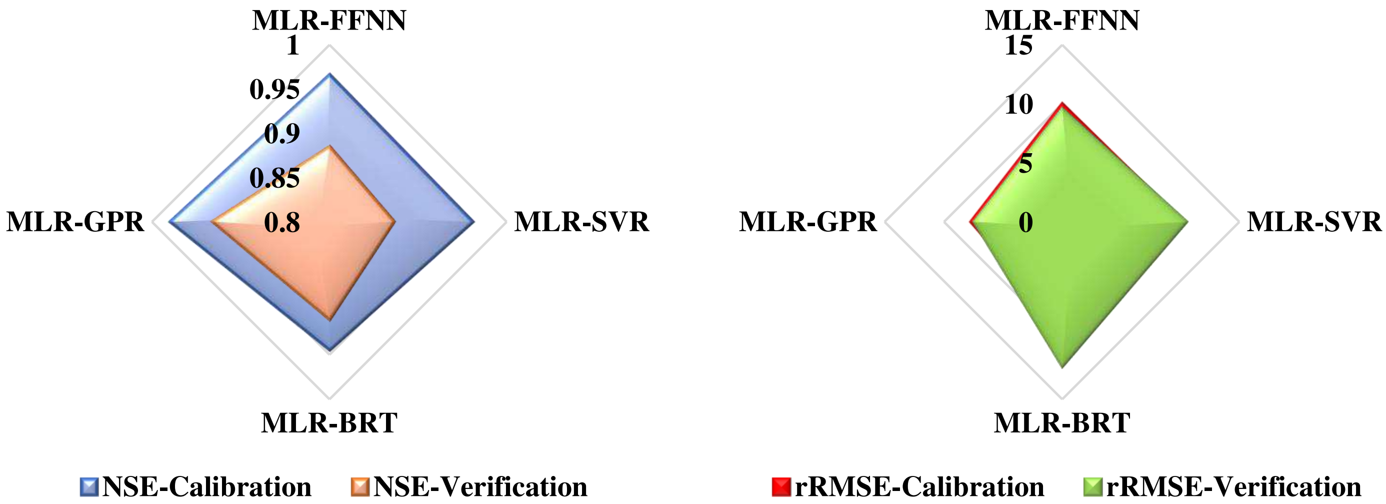


Figure 13: Radar Plot showing NSE, and rRMSE values for the hybrid models in calibration and verification stage

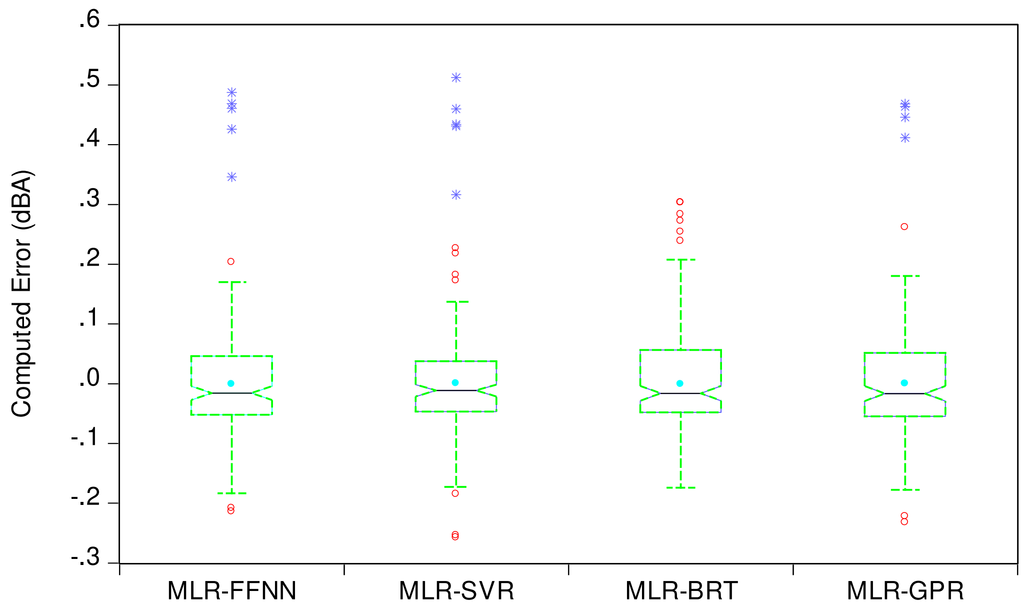


Figure 14: Boxplot for the absolute forecasted error in the prediction of the traffic noise by the hybrid models

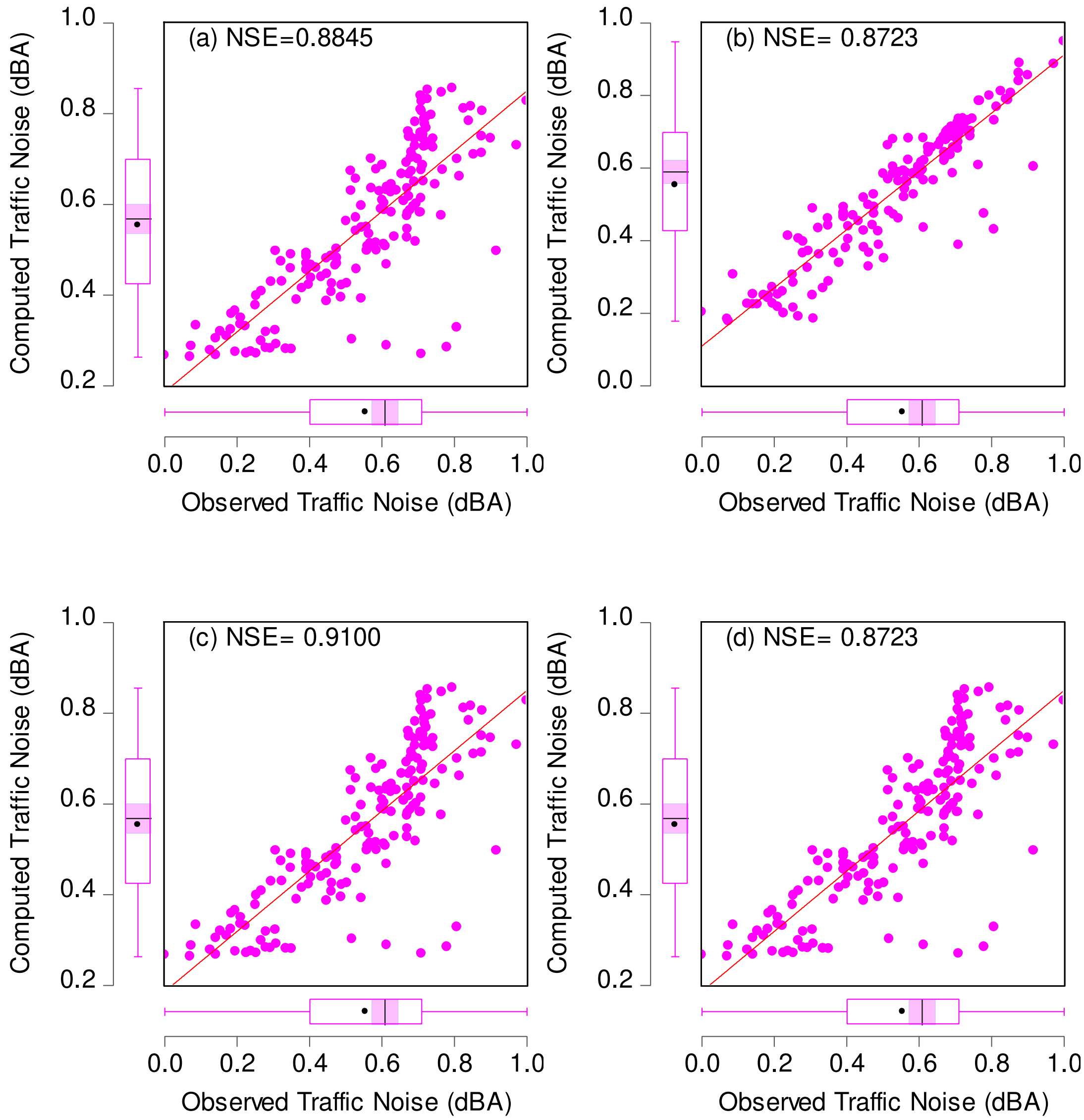


Figure 15: Scatter plots for the hybrid models (a) MLR-FFNN (a) MLR-SVR (c) MLR-BTR (d) MLR-GPR

5. Conclusions

Roadway traffic noise of Nicosia, North Cyprus was simulated by recording data from 12 different sites on four different road classes. The average traffic noise in the study area exceeds the safe noise level of 55 dBA recommended for European countries. The traffic noise was found to be higher along the expressway and least on local roads. The observed traffic noise levels ranged between 80.5 and 56.3 dBA with maximum value been recorded during the evening hours. The obtained data were used to predict the roadway traffic noise using five data driven models; one linear (MLR) and four nonlinear models (BRT, GPR, FFNN and SVR). Selection of relevant input variables prior to the model development, which is an essential step for obtaining appropriate performance in machine learning was done using MI and the result revealed that B, C, HV, MV and V are the most important parameters for the prediction of roadway traffic noise in Nicosia. The results of the models showed that the nonlinear models could predict the traffic noise with higher skills than the linear model with an improved performance of 20.9%, 16.9%, 10.33% and 12.64% compared to the linear model (MLR) for BRT, GPR, SVR and FFNN, respectively. Subsequently, four linear-nonlinear hybrid models were applied for improving the performance of the single models. Performance evaluation of the hybrid models using NSE, RMSE, MAE and rRMSE indicated that, the hybrid models demonstrated higher prediction capability than their equivalent single models with hierarchical order of MLR-GPR > MLR-BRT > MLR-FFNN > MLR-SVR. The MLR-GPR hybrid model improved the efficiency of the MLR and GPR models in the verification stage up to 27.26% and 10.30%, respectively. This shows that for the prediction of roadway traffic noise, stronger nonlinear models performed better when incorporated with linear models. The result of this study can be useful for the stakeholders in predicting noise level across the Nicosia which could further be used in developing noise maps with higher accuracy. The result of the study indicated that, inhabitants of the study area are exposed to an increased noise level of 14.5 dBA (on average) above the safe noise level (55 dBA). Therefore, immediate facilities to minimize the noise from the traffic are required for the wellbeing of the people along the roads. The effectiveness of other linear models (e.g., Autoregressive integrated mean average ARIMA, generalized linear regression, stepwise regression) for developing the linear-nonlinear hybrid models for estimation of roadway noise can be studied as future studies.

Conflict of Interest

The authors declare that they have no conflict of interest.

Data Availability Statement:

The Data used in this research is always available upon request through the editor.

References

Ahmed, A. A., & Pradhan, B. (2019). Vehicular traffic noise prediction and propagation modelling using neural networks and geospatial information system. *Environmental Monitoring and Assessment*, *191*(3), 190. https://doi.org/10.1007/s10661-019-7333-3

Athavale, J., Yoda, M., & Joshi, Y. (2019). Comparison of data driven modeling approaches for temperature prediction in data centers. *International Journal of Heat and Mass Transfer*, *135*, 1039–1052. https://doi.org/10.1016/j.ijheatmasstransfer.2019.02.041

Bonakdari, H., Ebtehaj, I., Samui, P., & Gharabaghi, B. (2019). Lake water-level fluctuations forecasting using Minimax Probability Machine Regression, Relevance Vector Machine, Gaussian Process Regression, and Extreme Learning Machine. *Water Resources Management*, *33*(11), 3965–3984. https://doi.org/10.1007/s11269-019-02346-0

Bravo-Moncayo, L., Lucio-Naranjo, J., Chávez, M., Pavón-García, I., & Garzón, C. (2019). A machine learning approach for traffic-noise annoyance assessment. *Applied Acoustics*, *156*, 262–270. https://doi.org/10.1016/j.apacoust.2019.07.010

Cai, H., Jia, X., Feng, J., Li, W., Hsu, Y. M., & Lee, J. (2020). Gaussian Process Regression for numerical wind speed prediction enhancement. *Renewable Energy*, *146*, 2112–2123. https://doi.org/10.1016/j.renene.2019.08.018

Chen, L., Liu, T., Tang, B., Xiang, H., & Sheng, Q. (2020). Modelling traffic noise in a wide gradient interval using artificial neural networks. *Environmental Technology*. https://doi.org/10.1080/09593330.2020.1734098

Covaciu, D., Florea, D., & Timar, J. (2015). Estimation of the noise level produced by road traffic in roundabouts. *Applied Acoustics*, *98*, 43–51. https://doi.org/10.1016/j.apacoust.2015.04.017

Doğan, E., & Akgüngör, A. P. (2013). Forecasting highway casualties under the effect of railway development policy in Turkey using artificial neural networks. *Neural Computing and Applications*, *22*(5), 869–877. https://doi.org/10.1007/s00521-011-0778-0

Ece, M., Tosun, I., Ekinci, K., & Yalçindağ, N. S. (2018). Modeling of road traffic noise and traffic flow measures to reduce noise exposure in Antalya metropolitan municipality.

Journal of Environmental Health Science and Engineering, *16*(1), 1–10. https://doi.org/10.1007/s40201-018-0288-4

Elith, J., Leathwick, J. R., & Hastie, T. (2008). A working guide to boosted regression trees. *Journal of Animal Ecology*, *77*(4), 802–813. https://doi.org/10.1111/j.1365-2656.2008.01390.x

European Environment Agency. (2014). *EEA Report No 10/2014 - Noise in Europe 2014*. kaduna. Retrieved from http://www.eea.europa.eu/publications/noise-in-europe-2014

Fan, J., Yue, W., Wu, L., Zhang, F., Cai, H., Wang, X., … Xiang, Y. (2018). Evaluation of SVM , ELM and four tree-based ensemble models for predicting daily reference evapotranspiration using limited meteorological data in di ff erent climates of China. *Agricultural and Forest Meteorology*, *263*, 225–241. https://doi.org/10.1016/j.agrformet.2018.08.019

Garg, N., & Maji, S. (2014). A critical review of principal traffic noise models: Strategies and implications. *Environmental Impact Assessment Review*, *46*, 68–81. https://doi.org/10.1016/j.eiar.2014.02.001

Gundogdu, O., Gokdag, M., & Yuksel, F. (2005). A traffic noise prediction method based on vehicle composition using genetic algorithms. *Applied Acoustics*, *66*, 799–809. https://doi.org/10.1016/j.apacoust.2004.11.003

Gökdag, M. (2012). Environmental health study of the road traffic noise in Erzurum-Turkey. *Iranian Journal of Environmental Health Science & Engineering*, *9*(22), 1–4. https://doi.org/http://www.ijehse.com/content/9/1/22

Hamad, K., Ali Khalil, M., & Shanableh, A. (2017). Modeling roadway traffic noise in a hot climate using artificial neural networks. *Transportation Research Part D: Transport and Environment*, *53*, 161–177. https://doi.org/10.1016/j.trd.2017.04.014

Hastie, T., Tibshirani, R., Friedman, J., 2011. The Elements of Statistical Learning. Springer.

Jahani, B., & Mohammadi, B. (2019). A comparison between the application of empirical and ANN methods for estimation of daily global solar radiation in Iran. *Theoretical and Applied Climatology*, *137*(1–2), 1257–1269. https://doi.org/10.1007/s00704-018-2666-3

Kim, S., & Singh, V. P. (2014). Modeling daily soil temperature using data-driven models and spatial distribution. *Theoretical and Applied Climatology*, *118*(3), 465–479. https://doi.org/10.1007/s00704-013-1065-z

Kumar, P., Nigam, S. P., & Kumar, N. (2014). Vehicular traffic noise modeling using artificial neural network approach. *Transportation Research Part C: Emerging Technologies*, *40*, 111–122. https://doi.org/10.1016/j.trc.2014.01.006

Müller, G., & Möser, M. (2013). *Handbook of engineering acoustics*. *Handbook of Engineering Acoustics*. Springer-Verlag Berlin Heidelberg. https://doi.org/10.1007/978-3-540-69460-1

Nedic, V., Despotovic, D., Cvetanovic, S., Despotovic, M., & Babic, S. (2014). Comparison of classical statistical methods and artificial neural network in traffic noise prediction. *Environmental Impact Assessment Review*, *49*, 24–30. https://doi.org/10.1016/j.eiar.2014.06.004

Nourani, V., Kisi, Ö., & Komasi, M. (2011). Two hybrid Artificial Intelligence approaches for modeling rainfall-runoff process. *Journal of Hydrology*, *402*, 41–59. https://doi.org/10.1016/j.jhydrol.2011.03.002

Nourani, V., & Fard-Sayyah, M. (2012). Sensitivity Analysis of the artificial neural network outputs in simulation of the evaporation process at different climatologic regimes. *Advances in Engineering Software*, *47*, 127–129.

Nourani, V., Khanghah, T. R., & Baghanam, A. H. (2015). Application of entropy concept for input selection of wavelet-ANN based rainfall-runoff modeling. *Journal of Environmental Informatics*, *26*(1), 52–70. https://doi.org/10.3808/jei.201500309

Nourani, V., Elkiran, G., Abdullahi, J., & Tahsin, A. (2019a). Multi-region modeling of daily global solar radiation with artificial intelligence ensemble. *Natural Resources Research*. https://doi.org/10.1007/s11053-018-09450-9

Nourani, V., Uzelaltinbulat, S., Sadikoglu, F., & Behfar, N. (2019b). Artificial intelligence based ensemble modeling for multi-station prediction of precipitation. *Atmosphere*, *10*(2), 80. https://doi.org/10.3390/atmos10020080

Nourani, V., Gökçekuş, H., & Umar, I. K. (2020a). Artificial intelligence based ensemble model for prediction of vehicular traffic noise. *Environmental Research*, *180*, 108852. https://doi.org/10.1016/j.envres.2019.108852

Nourani, V., Gökçeku, H., Umar, I. K., & Najafi, H. (2020b). An emotional artificial neural network for prediction of vehicular traffic noise. *Science of the Total Environment*, *707*, 136134. https://doi.org/10.1016/j.scitotenv.2019.136134

Ilgurel, N., Akdag, N. Y., & Akdag, A. (2016). Evaluation of noise exposure before and after noise barriers, a simulation study in Istanbul. *Journal of Environmental Engineering and Landscape Management*, *24*(4), 293–302. https://doi.org/10.3846/16486897.2012.721784

Rabehi, A., Guermoui, M., & Lalmi, D. (2020). Hybrid models for global solar radiation prediction: a case study. *International Journal of Ambient Energy*, *41*(1), 31–40. https://doi.org/10.1080/01430750.2018.1443498

Rasmussen, C. E. (2004). *Gaussian Processes in Machine Learning*. Springer-Verlag Berlin Heidelberg.

Sandberg, U., and Ejsmont, J. A. (2002). Tyre/Road Noise Reference Book. INFORMEX, Kisa, Sweden.

Schlittmeier, S., Feil, A., Liebl, A., & Hellbrück, J. (2015). The impact of road traffic noise on cognitive performance in attention-based tasks depends on noise level even within moderate-level ranges. *Noise and Health*, *17*(76), 148. https://doi.org/10.4103/1463-1741.155845

Sharma, A., Vijay, R., Bodhe, G. L., & Malik, L. G. (2018). An adaptive neuro-fuzzy interface system model for traffic classification and noise prediction. *Soft Computing*, *22*(6), 1891–1902. https://doi.org/10.1007/s00500-016-2444-z

Singh, D., Nigam, S. P., Agrawal, V. P., & Kumar, M. (2016). Vehicular traffic noise prediction using soft computing approach. *Journal of Environmental Management*, *183*, 59–66. https://doi.org/10.1016/j.jenvman.2016.08.053

Sørensen, M., Andersen, Z. J., Nordsborg, R. B., Becker, T., Tjønneland, A., Overvad, K., & Raaschou-Nielsen, O. (2013). Long-term exposure to road traffic noise and incident diabetes: a cohort study. *Environmental Health Perspectives*, *121*(2), 217–222.

https://doi.org/10.1289/ehp.1205503

Suleiman, A., Tight, M. R., & Quinn, A. D. (2016). Hybrid neural networks and boosted regression tree models for predicting roadside particulate matter. *Environmental Modeling and Assessment*, *21*(6), 731–750. https://doi.org/10.1007/s10666-016-9507-5

Taylor, K. E. (2001). Summarizing multiple aspects of model performance in a single diagram. *Journal of Geophysical Research*, *106*, 7183–7192. https://doi.org/http://dx.doi.org/10.1029/ 2000JD900719

Vapnik, V. N. (1998). *Statistical learning theory*. (S. Haykin, Ed.). New York: Wiley.

Vijay, R., Sharma, A., Chakrabarti, T., & Gupta, R. (2015). Assessment of honking impact on traffic noise in urban traffic environment of Nagpur, India. *Journal of Environmental Health Science and Engineering*, *13*(10), 1–10. https://doi.org/10.1186/s40201-015-0164-4

Wang, W., Xu, D., Chau, K., & Chen, S. (2015). Improved annual rainfall-runoff forecasting using PSO–SVM model based on EEMD. *Journal of Hydroinformatics*, *15*(4), 1377–1390. https://doi.org/10.2166/hydro.2013.134

Yaseen, Z. M., Deo, R. C., Hilal, A., Abd, A. M., Bueno, L. C., Salcedo-Sanz, S., & Nehdi, M. L. (2018). Predicting compressive strength of lightweight foamed concrete using extreme learning machine model. *Advances in Engineering Software*, *115*, 112–125. https://doi.org/10.1016/j.advengsoft.2017.09.004

Youssef, A. M., Pourghasemi, H. R., Pourtaghi, Z. S., & Al-Katheeri, M. M. (2016). Landslide susceptibility mapping using random forest, boosted regression tree, classification and regression tree, and general linear models and comparison of their performance at Wadi Tayyah Basin, Asir Region, Saudi Arabia. *Landslides*, *13*(5), 839–856. https://doi.org/10.1007/s10346-015-0614-1

Zhang, Y., Luo, L., Yang, J., Liu, D., Kong, R., & Feng, Y. (2019). A hybrid ARIMA-SVR approach for forecasting emergency patient flow. *Journal of Ambient Intelligence and Humanized Computing*, *10*(8), 3315–3323. https://doi.org/10.1007/s12652-018-1059-x