

## Research Article

# Sound Pressure Level at Terminals through Data Mining

**Luis Pastor Monteza** <sup>1</sup>, **Marco Castro Rivera** <sup>1</sup>, **Lenin Quiñones Huatangari** <sup>1</sup>,  
**and Eli Morales Rojas** <sup>1,2</sup>

<sup>1</sup>*Instituto de Investigación de Ciencia de Datos, Universidad Nacional de Jaén, Jaén, Peru*

<sup>2</sup>*Facultad de Ciencias Naturales y Aplicadas, Universidad Nacional Intercultural Fabiola Salazar Leguía de Bagua (UNIFSLB), Bagua 01721, Peru*

Correspondence should be addressed to Lenin Quiñones Huatangari; [lenin.quinones@unj.edu.pe](mailto:lenin.quinones@unj.edu.pe)

Received 13 April 2022; Revised 22 December 2022; Accepted 4 January 2023; Published 13 January 2023

Academic Editor: Diego Oliva

Copyright © 2023 Luis Pastor Monteza et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The land transportation is a cause of noise in cities, thus breaking the natural balance and bringing with it physiological and mental illnesses, as well as occupational accidents. In this sense, the objective of the research was to estimate the sound pressure in land terminals in the city of Jaen, Peru, using data mining algorithms. The methodology consisted in environmentally monitoring six terminals in the city of Jaen, during 2019, using a class 1 sound level meter; the exploratory analysis of the collected variables that influence the noise of the terminals (minimum and maximum sound pressure level, number of light and heavy vehicles, and equivalent sound pressure level) was performed, which were grouped into three groups of variables for the purpose of using data mining algorithms. Three algorithms were used, namely, artificial neural network, linear regression, and M5Rules, using the free software Weka. Considering all variables, the M5Rules method performed the best, because the value of the mean absolute error (0.7462), the root mean square error (1.0575), and uncertainty analysis (0.09) was the smallest compared to the other two methods. However, for the two remaining groups of variables, the linear regression model showed the lowest mean absolute error and mean square root of the error; in addition to presenting coefficients of determination close to one. The algorithms show good behavior when estimating the sound pressure of the terminals in the city of Jaen.

## 1. Introduction

Noise is a pollutant that is defined as the emission of energy originated by a set of aerial vibratory phenomena that, perceived by the auditory system, can cause discomfort or ear injuries [1]. Transportation, in addition to accelerated population growth, are the main causes of noise in cities, breaking the natural balance and harming people [2]. However, it causes negative effects on the health of others, bringing with it physiological and mental illnesses [3]. Noise also has implications for society, for example, impact on real estate prices and occupational accidents [4, 5]. Studies have correlated traffic noise variables with respiratory mortality [6]. Therefore, urban local governments should develop noise action plans; in this regard, geographic information on urban characteristics and noise levels is available in some places through smartphone applications [7, 8].

Vehicular traffic variables (flow and type of vehicle, etc.) explain a high variability in noise levels [9] and, in this sense, noise mapping models can be used. The application of Geo statistical methodology for the prediction and assessment of noise distribution in study areas made probability maps for noise levels, for control and monitoring in cities [10]. Simplification methods are becoming a real need for the acoustic community in developing countries, which allows to generate cost-effective traffic noise maps of a small city [11].

Models after the 1960's were not intended to predict single-vehicle levels but to predict the equivalent continuous level for traffic over a period. The commonly used models FHWA STAMINA, FHWA TNM, MITHRA, CoRTN, RLS90, STL-86, and ASJ-1993 have been analysed [12]. On the other hand, they compare the following models: the European standard, common noise assessment methods for EU member states (hereafter CNOSSOS), Nord2000, and the

traffic noise exposure (TRANEX) model based on the UK methodology, in terms of their source and propagation characteristics. The CNOSSOS, Nord2000, and TRANEX models are compared by analysing the estimated noise [13]. In the work of Álvarez Rodenbeek and Suárez Silva [14], they mention that road traffic becomes the main source of noise pollution in Chilean cities; therefore, they validated mathematical models such as SP-48 and SP-96, the Swiss STL-86, the English CoRTN, the one developed by CONAMA01, and the Sánchez model calibrated in Valdivia to predict road traffic noise.

They developed a noise prediction model using multiple regression analysis [15]. They employed machine learning modelling techniques to accurately predict road traffic noise; the techniques applied were decision tree regression, support vector machine, and artificial neural network (ANN) [16]. Artificial neural networks were applied for traffic noise prediction; the results show that the ANN approach is superior in predicting the traffic noise level to any other statistical methods [17]. In addition, they presented the application of emotional artificial neural network as a new generation neural network method for modelling road traffic noise [18]. To predict the noise level, an adaptive neurofuzzy inference system was developed and a detailed comparative analysis was performed with conventional soft-computing techniques such as ANNs, generalized linear model, random forests, decision trees, and support vector machines [19].

Public transportation in Peru is growing rapidly; in Lima alone, by 2019, it ranked seventh in the world traffic ranking with 57% congestion on its roads [20]. In Huancayo, Arana Velarde [21] determined that there is a relationship between the growth of urban mobility, the population increase, and the increase of the population. The effects fall on the population [22]. Therefore, it is necessary to model traffic noise for proper management in the city of Jaen; there are areas where motor vehicle activity is at a higher magnitude, such as bus stops, cars, and motorcycle cabs. These places generate noise levels that exceed those established in Peru's environmental quality standards [23, 24].

One of the relevant topics is the attention and design of activities based on science modelling and real situations [25]. Support vector machine (SVM), multivariate adaptive regression splines (MARS), and random forest (RF) techniques have been applied to estimate the densimetric Froude number approximation for the incipient movement of riprap stones [26]. Robust artificial intelligence models, namely, multivariate adaptive regression spline (MARS), gene-expression programming (GEP), and M5 model tree were employed to extract precise formulation for the pipes break rate estimation [27]. In this study, standardized precipitation evaporation rate values are formulated for various climates using three robust artificial intelligence models, namely, gene expression programming (GEP), model tree (MT), and multivariate adaptive regression spline (MARS) [28]. Models to estimate traffic noise are adequate tools to guide mitigation measures [29]. Although traffic noise models exist worldwide, they fail to generalize, because local conditions, e.g., vehicle type and climate vary from one location to another [30].

The research is justified because it is necessary to know the sound pressure level generated by the activity of land terminals in the city of Jaen, since it is a determinant of the quality of life of the inhabitants of the city, given the effects it has on their health and well-being.

In this context, the objective of the work was to estimate the sound pressure level of six land terminals located in the city of Jaen by means of data mining algorithms. The models obtained consider the conditions of the study area; these will be made available to institutions or persons interested in evaluating noise pollution and thus take appropriate actions to mitigate this problem.

## 2. Materials and Methods

The methodology consisted of four steps (Figure 1) as follows: (A) Environmental noise monitoring in six land terminals in the city of Jaen, Peru. Once the data were collected, they were cleaned and integrated. (B) Exploratory analysis of the data set to understand its nature and behavior. (C) Grouping of variables that influence the sound pressure level (LAeqt). (D) Use and evaluation of the behavior of data mining algorithms.

*2.1. Environmental Noise Monitoring.* Data were collected at six terminals in the city of Jaen (Figure 2) and characteristics of the study area were determined. Sampling was non-probabilistic, where the choice of land terminals was based on the vehicular flow and activities of each terminal. Sound pressure levels were measured during daytime hours, five days a week, from Thursday to Monday; it started on Monday, November 25 and ended on Sunday, December 22, 2019. The measurement schedule contemplated three shifts as follows: morning shift from 07:01 to 09:30 hours, midday shift from 12:00 to 15:00 hours, and afternoon shift from 18:30 to 21:00 hours.

Sol del Norte was located on Alfonso Villanueva Pinillos Street, in front of the Peru-Ecuador binational park, with exits for light vehicle exits; there was fluid vehicular traffic during the three measurement shifts. Malca was located between Pedro Cornejo Neyra and Alfredo Bastos streets, with exits for light vehicles and interprovincial buses; there was fluid vehicular traffic during the three measurement shifts, where cars, motorcycle cabs, and linear motorcycles were differentiated. Señor de Huamantanga was located on Mesones Muro Avenue, with interprovincial bus departures; there was fluid vehicular traffic in the morning and midday shifts; on the other hand, the afternoon shift was discarded because there were no bus departures, and the arrival of buses was variable, most of which were outside of the measurement schedule. Crucero Jaen and Turismo Fernández were located on Mesones Muro Avenue, with interprovincial bus departures; there was fluid vehicular traffic during the three measurement shifts, where cars, motorcycle cabs, linear motorcycles, and heavy vehicles were differentiated. Tetsur is located between Pedro Cornejo Neyra and La Marina streets, with interprovincial bus departures; there was fluid vehicular traffic in the midday and

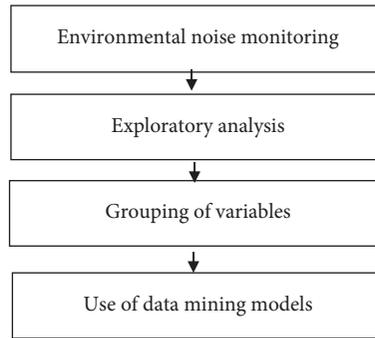


FIGURE 1: Methodology of the study consisting of four steps.

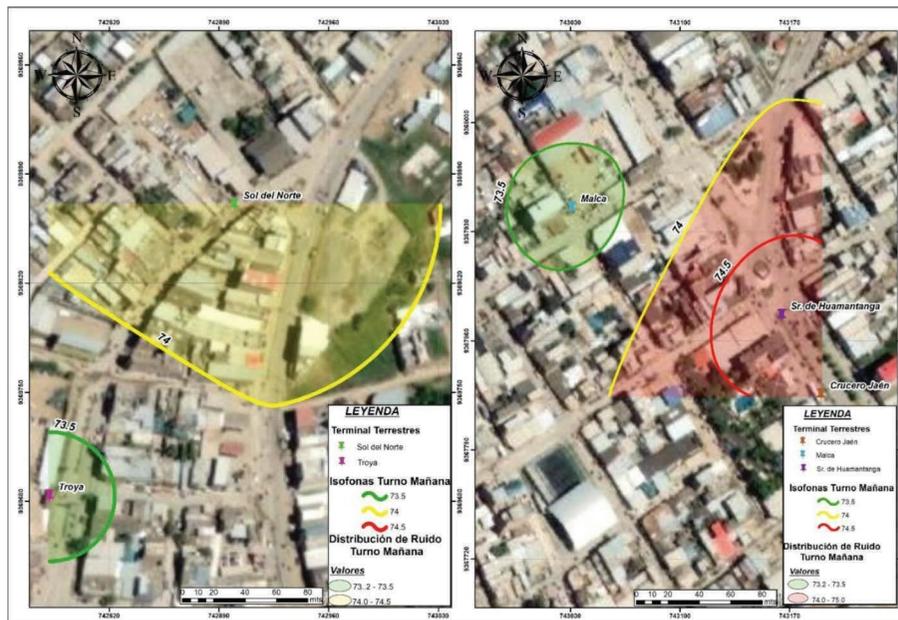


FIGURE 2: Location map of the study area.

afternoon shifts; however, the morning shift was discarded because activities began at 10:00 a.m., which was outside the measurement schedule; at this terminal, cars, motorcycle cabs, linear motorcycles, and heavy vehicles were differentiated. Troya was located on Mariscal Castilla Street, with exits for light vehicles; there was fluid vehicular traffic in the morning and midday shifts; in the afternoon shift, the measurement was not performed because activities ended at 18:00 hours; at this terminal, cars, motorcycle cabs, and linear motorcycles were differentiated.

Eight numerical variables were measured at the six land terminals in the city of Jaen as shown in Table 1.

**2.2. Grouping of Variables.** Of the seven independent variables described in Table 1, these have been grouped into three groups as follows: Scheme 1 composed of all variables. Scheme 2 made up of  $L_{min}$ ,  $L_{max}$ , motorcycle cabs, linear motorcycle, automobiles, and heavy vehicle. Finally, Scheme 3 is made up of  $L_{min}$ ,  $L_{max}$ , motorcycle cabs, linear motorcycle, and automobiles.

**2.3. Exploratory Data Analysis.** Table 2 shows the central tendency and variance statistics for the dependent variable LAeqt at the six land terminals.

**2.4. Data Mining Methods.** ANN is a computational model that emulates the biological neuronal system in information processing [31]. The multilayer perceptron ANN (MLP) is a type of network, in which its architecture consists of several layers of interconnected nodes or neurons, each of which is connected to all the neurons in the next layer. The input layer is where the data are presented to the neural network, while the output layer contains the response of the neural network [32]. Linear regressions are statistical models in which the value of the parameters is linearly proportional to the dependent variable [33]. The M5Rules algorithm generates a decision list for linear regression problems and uses the divide and conquer idea [34]. There are other models that were not addressed in this study, such as gene expression programming (GEP), a type of an evolutionary computation (EC) algorithm based on genotype/phenotype. It has also

TABLE 1: Name, description, and type of variables collected.

Variable	Description	Type
$L_{\min}$	Minimum sound pressure level, obtained in each repetition of 1 minute	Continue
$L_{\max}$	Maximum sound pressure level, obtained in each repetition of 1 minute	Continue
Motorcycle cabs	Number of motorcycles with three wheels and roof, categorized as light vehicles	Discrete
Linear motorcycle	Number of two-wheeled motorcycles, categorized as light vehicles	Discrete
Automobiles	Number of motorized vehicles consisting of four wheels, categorized as light vehicles	Discrete
Heavy vehicle	Number of transport vehicles with four or more wheels and a load capacity of more than 3.5 tons	Discrete
Buses	Number of motor vehicle with capacity for a large number of passengers, intended for passenger transportation	Discrete
$LA_{eqT}$	Equivalent continuous sound pressure level, in a period $T$ , A-weighted, measured by a class 1 sound level meter	Continue

TABLE 2: Measures of central tendency and variance of  $LA_{eqT}$ .

Statistic	Sol norte	Troya	Malca	Fernández	Señor de huamantanga	Tetsur
Minimum	72.6	71.65	71.07	73.47	73.6	72.85
Median	74.15	73.45	72.6	74.52	75.22	73.85
Mean	74.15	73.37	72.72	74.55	75.08	73.95
Maximum	75.31	74.65	74.5	76.2	77.1	75.25
Standard deviation	0.766	0.879	1.007	0.644	0.865	0.714
Variance	0.586	0.773	1.015	0.414	0.749	0.51

been used in petroleum engineering for viscosity estimation [35, 36]. The multivariate adaptive regression splines (MARS) analysis model, unlike the (GEP) model, is little used in health issues; it is a nonparametric modelling method developed systematically and automatically without limitation of the assumptions that traditional regression models must meet [37].

2.5. *Evaluation of Model Performance.* It was carried out through cross-validation, using the mean absolute error tests (MAE), the root mean square of the error (RMSE), and the coefficient of determination ( $R^2$ ). These values have been calculated by means of the equations (1)–(3) [38].

$$MAE = \frac{\sum_{i=1}^n |Q_i^{\text{mod}} - Q_i^{\text{obs}}|}{n}, \quad (1)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (Q_i^{\text{mod}} - Q_i^{\text{obs}})^2}{n}}, \quad (2)$$

$$R^2 = \left[ \frac{\sum_{i=1}^n (Q_i^{\text{obs}} - Q^{-\text{obs}})(Q_i^{\text{mod}} - Q^{-\text{mod}})}{\sqrt{\left[ \sum_{i=1}^n (Q_i^{\text{obs}} - Q^{-\text{obs}})^2 \right]} \sqrt{\left[ \sum_{i=1}^n (Q_i^{\text{mod}} - Q^{-\text{mod}})^2 \right]}} \right]^2, \quad (3)$$

where  $Q_i^{\text{mod}}$  is the  $i$ -th value produced by the model,  $Q_i^{\text{obs}}$  is the  $i$ -th value produced by measurement,  $n$  is the number of observations,  $Q^{-\text{obs}}$  is the mean of the observed values, and  $Q^{-\text{mod}}$  is the mean of the values produced by the model.

2.6. *Performance of Uncertainty and Reliability.* The objective of uncertainty analysis is to restrict the expected range in which the true value of the result of an

experiment lies. This estimated range is in the form of an interval. One uncertainty analysis procedure is  $U95$ . The value of  $U95$  associated with the result of a given experiment can be interpreted; by performing the given experiment repeatedly over and over again, the true value of the result of that experiment will lie in the offered uncertainty range for approximately 95 times out of 100 trials. Moreover, the value of  $U95$  is given by the following equation [39]:

TABLE 3: Behavior of estimated models LAeqt.

Model	Scheme 1					Scheme 2					Scheme 3				
	RMSE	$R^2$	MAE	U95	Rel	RMSE	$R^2$	MAE	U95	Rel	RMSE	$R^2$	MAE	U95	Rel
MLP	1.60	0.68	1.26	0.10	100	1.70	0.63	1.35	0.10	99.89	1.69	0.63	1.34	0.10	99.86
Linear regression	1.35	0.76	1.11	0.09	100	1.48	0.71	1.17	0.10	99.96	1.18	0.71	1.48	0.10	99.93
M5Rules	1.40	0.75	1.06	0.09	99.93	1.57	0.68	1.14	0.10	99.90	1.57	0.67	1.14	0.10	99.86

$$U95 = \left(\frac{1.96}{n}\right) \sqrt{\sum_{i=1}^n (Q_i^{\text{obs}} - Q^{-\text{obs}})^2 + \sum_{i=1}^n (Q_i^{\text{obs}} - Q_i^{\text{mod}})^2}. \quad (4)$$

The reliability analysis is a statistical method for measuring the overall consistency of a model. It determines if a suggested model achieves a permissible level of performance. In general, a metric for the reliability analysis is defined by the following equation [39].

$$\text{Reliability} = \left(\frac{100\%}{n}\right) \sum_{i=1}^n k_i, \quad (5)$$

where  $k_i$  is obtained through two steps. First, the relative average error (RAE) is defined as a vector whose  $i$ -th component is given by the following equation.

$$\text{RAE}_i = \left| \frac{Q_i^{\text{obs}} - Q_i^{\text{mod}}}{Q_i^{\text{obs}}} \right|, \quad (6)$$

Next, if  $\text{RAE}_i \leq \Delta$ , then  $k_i = 1$ ; otherwise,  $k_i = 0$ , which is the threshold value of LAeqt parameter. In other words,  $k_i$  is

$$\text{LAeqt} = 0.052L \min + 0.002L \max + 0.306M + 0.222ML + 0.243A + 0.309VP + 63.943, \quad (7)$$

$$\text{LAeqt} = 0.054L \min + 0.002L \max + 0.306M + 0.224ML + 0.243A + 63.2646. \quad (8)$$

where  $M$  is the number of motorcycle cabs;  $ML$  is the number of linear motorcycles;  $A$  is the number of automobiles and  $VP$  is the number of heavy vehicles.

#### 4. Discussion

It is important to consider noise pollution in order to achieve sustainable cities [40]. However, sound pressure levels derived from vehicles in urban environments are becoming more pressing in third world cities, due to the lack of knowledge and low interest of the public administration in this problem [41]. The most immediate solution lies in reducing the flow of public transportation vehicles, as well as updating the vehicle fleet [42]. Vehicle reorganization can contribute to the improvement of air quality and the health of the resident population, as well as to the reduction of environmental pollution [43]. Likewise, in Peru, the Peruvian Congress has implemented strategies to minimize air pollution problems. Through the Organic Law of

defined as the number of times the value of RAE is less than or equal to that of  $\Delta$ . The optimum value of  $\Delta$  based on Chinese Standards is 0.2 or equivalently is 20%.

#### 3. Results

For scheme 1, the algorithm that showed the best performance in terms of the statistics RMSE,  $R^2$ , MAE was the M5Rules. However, for schemes 2 and 3, the linear regression model showed the best performance in estimating sound pressure. In addition, the U95 and reliability indices were calculated (Table 3). Table 3 shows that for scheme 1, the linear regression and M5rules algorithms presented the lowest value of uncertainty with a confidence level of 95% (0.09). In addition, ANN and linear regression for scheme 1 were at a higher level of reliability (100%).

The model obtained for scheme 1, using the M5Rules algorithm consisted of five rules, as shown in Table 4. In addition, for schemes 2 and 3, two multiple linear regression type models were obtained, as shown in equations (7) and (8), respectively.

Municipalities (Ley 27972), it was agreed to monitor signaling systems, land terminals, and pedestrian traffic [44].

Linear regression models can indicate predictive accuracy for any environmental factor [45]. By means of simple linear regression models as well as multiple linear regression models, they were adjusted to the conditions of the urban area and depend on the number of vehicles and equivalent noise [46]. In another study, the mean square error shows that the proposed regression model predicts sound pressure accurately [47]. This study shows that the estimation of sound pressure (0.7462) was efficient.

Data mining has different applications, such as the estimation of the behavior of chemical elements [48, 49] allows to determine predictive models for decision-makers [50], influencing factors of the variables that allow deciphering the behavior of the three models. Therefore, the noise characteristics generated by the vehicle fleet need to be known through economic techniques, such as data mining [51]. Data mining algorithms make it possible to identify

TABLE 4: Rules for determining LAeqT using the algorithm M5Rules.

Rule	Description
1	Si ( $M > 14.5$ y $ML > 5.5$ ) entonces $LAeqT = 0.0003L \min + 0.1569M + 0.0973ML + 0.2139A + 0.2675VP + 1.108B + 70.5385$
2	Si ( $M > 12.5$ y $A \leq 3.5$ ) entonces $LAeqT = 0.0495L \min + 0.3375M + 0.3342ML + 0.3255A + 0.9288VP + 1.5518B + 62.1908$
3	Si ( $M > 12.5$ ) entonces $LAeqT = 0.0018L \min + 0.001L \max + 0.236M + 0.2085ML + 0.1629A - 0.387VP + 1.3247B + 68.3128$
4	Si ( $ML \leq 6.5$ ) entonces $LAeqT = 0.0955L \min + 0.0235L \max + 0.3401M + 0.2417ML + 0.3008A + 1.1535VP + 2.5167B + 57.1125$
5	$LAeqT = 0.0382L \max + 0.4786M + 0.2522ML + 0.3821A + 61.1762$

urban reality, for example, the collective behavior of citizens. These results allow solving urban problems in transportation, environment, and public safety [52]. The Random Forest algorithm allows the elimination of data and obtain values adjusted to reality [53]. In addition, the Random Forest model was the most appropriate, a model that is widely used to estimate noise, and has been used in telecommunications, the activation of circuits with simple voice recognition, among others [54].

The tightest mean absolute error was for scheme one. The  $R^2$  for scheme one was 0.8310; the values were higher than the noise levels measured in London ( $R^2$ : 0.56–0.73) [55]. In this sense, the MAE method was used to validate the accuracy of the correlation and to compare the sound pressure level [56]. The MAPE yielded the lowest error on all the data evaluated, as stated by [49]. Also, the MAE shows the difference between modeled and observed values in absolute terms [57].

## 5. Conclusions

The estimation of sound pressure in six ground terminals was located in the city of Jaen using data mining algorithms. The M5Rules method was the one that presented the best performance, considering all the variables collected, due to the fact that a mean absolute error value of 0.7462 was obtained for the artificial neural networks and multiple linear regression algorithms. However, for scheme 2 and 3, the linear regression model was the one that showed the lowest mean absolute error and determination coefficients close to one, which presents a better behavior. It is concluded that data mining techniques could be used to estimate the sound pressure of the terminals for the city of Jaen, Peru.

## Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## Acknowledgments

The authors would like to thank the Professional School of Forestry and Environmental Engineering of the National University of Jaen for the guidance given to the researchers in the development of their research for their professional degree.

## References

- [1] J. Morales and J. y Fernández, “Análisis discriminante de algunas variables que influyen en la contaminación acústica debida al tráfico urbano en una gran ciudad,” *Revista Ingenierías Universidad de Medellín*, vol. 11, no. 21, pp. 13–22, 2012.
- [2] Y. Liu, X. Ma, L. Shu et al., “Internet of things for noise mapping in smart cities: state of the art and future directions,” *IEEE Network*, vol. 34, no. 4, pp. 112–118, 2020.
- [3] J. R. Quintero, “Caracterización del ruido producido por el tráfico vehicular en el centro de la ciudad de Tunja, Colombia,” *Revista Virtual Universidad Católica del Norte*, vol. 1, no. 36, pp. 311–343, 2012.
- [4] A. Tobías, J. Díaz, M. Saez, and J. C. Alberdi, “Use of Poisson regression and Box–Jenkins models to evaluate the short-term effects of environmental noise levels on daily emergency admissions in Madrid, Spain,” *European Journal of Epidemiology*, vol. 17, no. 8, pp. 765–771, 2001.
- [5] M. Picard, S. A. Girard, M. Simard, R. Larocque, T. Leroux, and F. Turcotte, “Association of work-related accidents with noise exposure in the workplace and noise-induced hearing loss based on the experience of some 240, 000 person-years of observation,” *Accident Analysis and Prevention*, vol. 40, no. 5, pp. 1644–1652, 2008.
- [6] A. Tobías, A. Recio, J. Díaz, and C. Linares, “Does traffic noise influence respiratory mortality?” *European Respiratory Journal*, vol. 44, no. 3, pp. 797–799, 2014.
- [7] P. Aumond, A. Can, G. Rey Gozalo, N. Fortin, and E. Suárez, “Method for in situ acoustic calibration of smartphone-based sound measurement applications,” *Applied Acoustics*, vol. 166, Article ID 107337, 2020.
- [8] B. Devkota, H. Miyazaki, A. Witayangkurn, and S. M. Kim, “Using volunteered geographic information and nighttime light remote sensing data to identify tourism areas of interest,” *Sustainability*, vol. 11, no. 17, p. 4718, 2019.
- [9] A. J. Torija, N. Genaro, D. P. Ruiz, A. Ramos-Ridao, M. Zamorano, and I. Requena, “Priorization of acoustic variables: environmental decision support for the physical characterization of urban sound environments,” *Building and Environment*, vol. 45, no. 6, pp. 1477–1489, 2010.
- [10] D. L. Yepes, M. Gómez, L. Sánchez, and A. C. y Jaramillo, “Metodología de elaboración de mapas acústicos como herramienta de gestión del ruido urbano-Caso Medellín,” *Dyna*, vol. 76, no. 158, pp. 29–40, 2009.
- [11] N. A. Bastián-Monarca, E. Suárez, and J. P. Arenas, “Assessment of methods for simplified traffic noise mapping of small cities: casework of the city of Valdivia, Chile,” *Science of the Total Environment*, vol. 550, pp. 439–448, 2016.
- [12] C. Steele, “A critical review of some traffic noise prediction models,” *Applied Acoustics*, vol. 62, no. 3, pp. 271–287, 2001.
- [13] J. Khan, M. Ketzler, S. S. Jensen, J. Gulliver, E. Thysell, and O. Hertel, “Comparison of road traffic noise prediction

- models: CNOSSOS-EU, Nord2000 and TRANEX,” *Environmental Pollution*, vol. 270, Article ID 116240, 2021.
- [14] J. P. Álvarez Rodenbeek and E. Suárez Silva, “Estudio comparativo para modelos predictivos del ruido de tráfico rodado, a través de mediciones in situ en un sector de la ciudad de Osorno,” *Síntesis Tecnológica*, vol. 4, no. 2, pp. 17–24, 2011.
- [15] C. Thakre, V. Laxmi, R. Vijay, D. J. Killedar, and R. Kumar, “Traffic noise prediction model of an Indian road: an increased scenario of vehicles and honking,” *Environmental Science and Pollution Research*, vol. 27, no. 30, pp. 38311–38320, 2020.
- [16] M. Ali Khalil, K. Hamad, and A. Shanableh, “Developing machine learning models to predict roadway traffic noise: an opportunity to escape conventional techniques,” *Transportation Research Record*, vol. 2673, no. 4, pp. 158–172, 2019.
- [17] V. Nedic, D. Despotovic, S. Cvetanovic, M. Despotovic, and S. Babic, “Comparison of classical statistical methods and artificial neural network in traffic noise prediction,” *Environmental Impact Assessment Review*, vol. 49, pp. 24–30, 2014.
- [18] V. Nourani, H. Gökçekuş, I. K. Umar, and H. Najafi, “An emotional artificial neural network for prediction of vehicular traffic noise,” *Science of the Total Environment*, vol. 707, Article ID 136134, 2020.
- [19] D. Singh, R. Upadhyay, H. S. Pannu, and D. Leray, “Development of an adaptive neuro fuzzy inference system based vehicular traffic noise prediction model,” *Journal of Ambient Intelligence and Humanized Computing*, vol. 12, no. 2, pp. 2685–2701, 2021.
- [20] Tom, “As Our World Changes, Traffic Tells the story,” 2022, [https://www.tomtom.com/en\\_gb/traffic-index/](https://www.tomtom.com/en_gb/traffic-index/).
- [21] F. Arana Velarde, “Incidencia de la movilidad urbana en el crecimiento insostenible de la ciudad de Huancayo,” *Arquitectura Y Urbanismo*, vol. 42, no. 3, pp. 112–117, 2021.
- [22] M.-P. Thomas, H. Aldy, L. Pattaroni, V. Kaufmann, and S. Galloux, *Choix résidentiels et modes de vie dans l’agglomération franco-valdo-Genévoise*, vol. 111, EPFL, Switzerland, 2011, [http://infoscience.epfl.ch/record/168665/files/Rapportfinal\\_FVG\\_Lasur\\_Modesdevie\\_2011\\_CdL3.pdf](http://infoscience.epfl.ch/record/168665/files/Rapportfinal_FVG_Lasur_Modesdevie_2011_CdL3.pdf).
- [23] C. K. Cruzado and Y. S. y Soto, *Evaluation of vehicular noise pollution based on Supreme Decree N 085-2003-PCM Regulation of Environmental Quality Standards for Noise carried out in the province of Jaén, department of Cajamarca*, Facultad de Ingeniería y Arquitectura, Universidad Peruana Unión, Tarapoto, Perú, 2017.
- [24] F. N. Silva, *Estándares de calidad ambiental (ECAS) para ruido en los Principales centros de Educación Superior Universitaria, de la ciudad de Jaén. Tesis de Licenciatura. Carrera Profesional de Ingeniería Forestal Ambiental*, Universidad Nacional de Jaén, Jaén, Perú, 2019.
- [25] A. Abassian, F. Safi, S. Bush, and J. Bostic, “Five different perspectives on mathematical modeling in mathematics education,” *Investigations in Mathematics Learning*, vol. 12, no. 1, pp. 53–65, 2020.
- [26] M. Najafzadeh and G. Oliveto, “Riprap incipient motion for overtopping flows with machine learning models,” *Journal of Hydroinformatics*, vol. 22, no. 4, pp. 749–767, 2020.
- [27] Y. Amiri-Ardakani and M. Najafzadeh, “Pipe break rate assessment while considering physical and operational factors: a methodology based on global positioning system and data-driven techniques,” *Water Resources Management*, vol. 35, no. 11, pp. 3703–3720, 2021.
- [28] A. Barzkar, M. Najafzadeh, and F. Homaei, “Evaluation of drought events in various climatic conditions using data-driven models and a reliability-based probabilistic model,” *Natural Hazards*, vol. 110, no. 3, pp. 1931–1952, 2022.
- [29] R. A. Melo, R. L. Pimentel, D. M. Lacerda, and W. M. Silva, “Applicability of models to estimate traffic noise for urban roads,” *Journal of Environmental Health Science and Engineering*, vol. 13, no. 1, pp. 83–187, 2015.
- [30] K. Hamad, M. Ali Khalil, and A. Shanableh, “Modeling roadway traffic noise in a hot climate using artificial neural networks,” *Transportation Research Part D: Transport and Environment*, vol. 53, pp. 161–177, 2017.
- [31] L. Quiñones Huatangari, L. Ochoa Toledo, N. Kemper Valverde, O. Gamarra Torres, J. Bazán Correa, and J. Delgado Soto, “Red neuronal artificial para estimar un índice de calidad de agua,” *Enfoque UTE*, vol. 11, no. 2, pp. 109–120, 2020.
- [32] R. Flórez and J. M. y Fernández, *Las Redes Neuronales Artificiales: Fundamentos Teóricos Y Aplicaciones Prácticas*, Netbiblo, La Coruña, España, 2008.
- [33] L. Palacios-Cruz, M. Pérez, R. Rivas-Ruiz, and J. O. Talavera, “From clinical judgment to linear regression model,” *Revista Médica del Instituto Mexicano del Seguro Social*, vol. 51, no. 6, pp. 656–661, 2013.
- [34] N. N. R. Ramírez, D. M. P. Ibarra, Á. R. Chávez, and I. A. M. y Bautista, “Sistema de diagnóstico de enfermedades de vías urinarias,” *Pistas Educativas*, vol. 36, no. 114, 2018.
- [35] Q. Li, W. Wang, S. Han, and J. Li, “Evolving classifier ensemble with gene expression programming,” *Third International Conference on Natural Computation (ICNC 2007)*, vol. 3, no. 1, pp. 546–550, 2007.
- [36] A. Rostami, A. Hemmati-Sarapardeh, and A. H. Mohammadi, “Estimating n-tetradecane/bitumen mixture viscosity in solvent-assisted oil recovery process using GEP and GMDH modeling approaches,” *Petroleum Science and Technology*, vol. 37, no. 14, pp. 1640–1647, 2019.
- [37] J. Vanegas and F. Vásquez, “Multivariate Adaptive Regression Splines (MARS), an alternative for the analysis of time series,” *Gaceta Sanitaria*, vol. 31, no. 3, pp. 235–237, 2017.
- [38] K. P. Singh, A. Basant, A. Malik, and G. Jain, “Artificial neural network modeling of the river water quality - a case study,” *Ecological Modelling*, vol. 220, no. 6, pp. 888–895, 2009.
- [39] F. Saberi-Movahed, M. Najafzadeh, and A. Mehrpooya, “Receiving more accurate predictions for longitudinal dispersion coefficients in water pipelines: training group method of data handling using extreme learning machine conceptions,” *Water Resources Management*, vol. 34, no. 2, pp. 529–561, 2020.
- [40] G. Rey Gozalo, E. Suárez, A. L. Montenegro, J. P. Arenas, J. M. Barrigón Morillas, and D. Montes González, “Noise estimation using road and urban features,” *Sustainability*, vol. 12, no. 21, p. 9217, 2020.
- [41] E. Gaja, A. Gimenez, S. Sancho, and A. Reig, “Sampling techniques for the estimation of the annual equivalent noise level under urban traffic conditions,” *Applied Acoustics*, vol. 64, no. 1, pp. 43–53, 2003.
- [42] A. Ramírez, E. A. Domínguez, and I. y Borrero, “El ruido vehicular urbano y su relación con medidas de restricción del flujo de automóviles,” *Revista de la Academia Colombiana de Ciencias Exactas, Físicas y Naturales*, vol. 35, no. 135, pp. 143–156, 2011.
- [43] V. Tapia, L. Carbajal, V. Vásquez et al., “Reordenamiento vehicular y contaminación ambiental por material particulado (2, 5 y 10), dióxido de azufre y dióxido de nitrógeno en Lima Metropolitana, Perú,” *Revista Peruana de Medicina Experimental y Salud Pública*, vol. 35, no. 2, pp. 190–197, 2018.

- [44] C. del Perú, *Ley orgánica de Municipalidades, Ley N° 27972. Congreso de la República del Perú. Diario Oficial El Peruano, Perú. 6 de mayo de 2003*, Editora Perú, Peru, 2003.
- [45] H. Tang, Y. Ding, and B. Singer, “Interactions and comprehensive effect of indoor environmental quality factors on occupant satisfaction,” *Building and Environment*, vol. 167, Article ID 106462, 2020.
- [46] K. V. Coral Carrillo, “Modelos estadísticos de ruido ambiental para el Distrito Metropolitano de Quito DMQ, mediante datos históricos del 2009 al 2015, como herramienta de calidad ambiental,” *ACI Avances en Ciencias e Ingenierías*, vol. 12, no. 1, p. 24, 2020.
- [47] M. Zhang, R. Hu, and L. Jiang, “Three-dimensional sound reproduction in vehicle based on data mining technique,” *Concurrency and Computation: Practice and Experience*, vol. 31, no. 4, 2019.
- [48] D. Abooli, R. Soleimani, and S. Gholamreza-Ravi, “Characterization of physico-chemical properties of biodiesel components using smart data mining approaches,” *Fuel*, vol. 266, Article ID 117075, 2020.
- [49] J. D. Martínez-Morales, J. M. Romo-Orozco, and E. R. y Palacios-Hernández, “Comparación de desempeño de las redes neuronales artificiales RBF y MLP para modelar el ruido en un área urbana,” *Memorias. XVI Congreso Latinoamericano de Control Automático. Cancún, México*. vol. 14-27, pp. 158–163, 2014.
- [50] M. He, L. R. e Sousa, T. Miranda, and G. Zhu, “Rockburst laboratory tests database—application of data mining techniques,” *Engineering Geology*, vol. 185, pp. 116–130, 2015.
- [51] E. Uludamar, E. Tosun, and K. Aydın, “Experimental and regression analysis of noise and vibration of a compression ignition engine fuelled with various biodiesels,” *Fuel*, vol. 177, pp. 326–333, 2016.
- [52] S. Wang, Y. Zhang, C. Zhang, and M. Yang, “Improved artificial neural network method for predicting photovoltaic output performance,” *Global Energy Interconnection*, vol. 3, no. 6, pp. 553–561, 2020.
- [53] H. Alaiz-Moreton, L. Fernández-Robles, J. Alfonso-Cendón, M. Castejón-Limas, L. Sánchez-González, and H. Pérez, “Data mining techniques for the estimation of variables in health-related noisy data,” *Memorias. International Workshop on Soft Computing Models in Industrial and Environmental Applications. León, España*, vol. 8, pp. 482–491, 2017.
- [54] R. F. Medina-Merino and C. I. Ñique-Chacón, “Bosques aleatorios como extensión de los árboles de clasificación con los programas R y Python,” *Interfases*, vol. 12, no. 10, pp. 165–189, 2017.
- [55] J. Gulliver, D. Morley, D. Vienneau et al., “Development of an open-source road traffic noise model for exposure assessment,” *Environmental Modelling and Software*, vol. 74, pp. 183–193, 2015.
- [56] M. Farzaneh-Gord, A. Arabkoohsar, and R. N. Koury, “Novel natural gas molecular weight calculator equation as a functional of only temperature, pressure and sound speed,” *Journal of Natural Gas Science and Engineering*, vol. 30, pp. 195–204, 2016.
- [57] C. L. Zuluaga, *Un aporte a la gestión del ruido urbano en colombia, caso de estudio municipio de Envigado. Tesis de Maestría. Escuela de Geociencias y Medio Ambiente, Universidad Nacional de Colombia, Medellín, Colombia*, 2009.