Journal of ENVIRONMENTAL Treatment Techniques

J. Environ. Treat. Tech.
ISSN: 2309-1185

Journal web link: http://www.jett.dormaj.com
https://doi.org/10.47277/JETT/10(3)227

Machine Learning Model for Reduction of Airborne Infections and Cognitive Load in a Car Cabin

Bankapalli Vamsi [1]*, Aasritha Bommadevara [2], Vanam Nagasri [3], Tejavath Tharun [3]

[1] Department of Production Engineering, National Institute of Technology, Tiruchirappalli, India
[2] Department of Instrumentation and control engineering, National Institute of Technology, Tiruchirappalli, India
[3] Department of Production Engineering, National Institute of Technology, Tiruchirappalli, India

Received: 06/08/2022 Accepted: 09/09/2022 Published: 20/09/2022

Abstract

The importance of medical care has grown massively, making it one of life's most essential components. People often use their vehicles in recirculation mode to provide optimum cooling in many cities with high air humidity and temperatures. On the other side, the recirculation mode of the cabin's air prevents O_2 from entering and causes a rise in CO_2. Increased health concerns, a decline in focus, and poor performance are all related to the increased CO_2 concentration brought on by human exhale and metabolism. The paper describes an experimental investigation on how carbon dioxide builds up in a car's interior as a result of metabolism and breathing by passengers; specific levels of this gas may be dangerous for everyone inside, especially drivers. It is critical to maintain cabin concentration levels within the authorized limits since inhaling this gas may impair a driver's ability to make intelligent decisions. Opening the cabin windows may be a practical, easy, and affordable way to do this. Opening the cabin windows, though might always make it less comfortable inside. As a response, given model passengers can temporarily open the windows may significantly affect how much CO_2 is present within the cabin. Using Our MVPR machine learning model and we demonstrate a GUI Framework, that can predict the forecast of CO_2 concentrations in a cabin at a particular time, temperature, and relative humidity to avoid negative health impacts caused by CO_2 gas.

keywords: Machine Learning, Airborne Infections, Occupational Health

1. Introduction

Automobile occupants in nations with "composite climates" such as India, utilize air conditioners to achieve thermal comfort. This can encourage automotive occupants to restrict the flow of outdoor air into the cabins, leading to a rise in CO_2 concentrations generated by the metabolic respiration of automobile occupants. This CO_2 accumulation contributes to health complications like "airborne infections" and "cognitive decision-making." Considering the recent SARS-Cov-2 (COVID-19) outbreak, it has become clear that sharing confined places with insufficient ventilation can facilitate the propagation of the virus (1). Furthermore, some research indicates that increased CO_2 concentrations in interior environments, without corresponding changes in ventilation rate, are negative to occupants' decision-making performance (2)(3). Consequently, it is essential to know the cabin concentration levels to make critical decisions (such as opening windows for a time) when traveling. In response, we attempted to develop a machine learning (ML) model that could predict the CO_2 concentrations in a cabin at the given time, temperature, and relative humidity.

2. Data collection

To formulate a model for prediction purposes, a training data set is required. This dataset is obtained by monitoring indoor air quality (IAQ) parameters such as CO_2 concentrations, temperature, and relative humidity. This monitoring is done with the help of commercial gas sensors by a Testo 400 monitor. During December in Jodhpur, Rajasthan, India, the monitoring is performed when there are four occupants inside the vehicle. The data was monitored once the monitoring device was installed and activated in the car. At 16:00 on December 17, 2021, all sensors were installed. After 16:20, observations of data recording were evaluated. The average CO_2 sensor measurement at 16:20 was 2168 ppm, indicating that data logging began at this concentration. After monitoring, all the data was imported into our workbooks for further analysis.

3. Model creation

An ML model for prediction was developed using a Non-linear or curvilinear regression methodology. In our study, we utilized the second-degree multivariate polynomial regression (MVPR) technique. In this technique, the relationship between independent and dependent variables is quadratic. In our case, there are three independent variables (time, temperature, and

***Corresponding author:** Bankapalli Vamsi, Department of Production Engineering, National Institute of Technology, Tiruchirappalli, E-mail: vamsibankapalli@gmail.com

relative humidity) and CO_2 concentration as the dependent variable. Therefore, the resulting general equation is –

$$Y = [\Theta_0 + e] + \begin{bmatrix}\Theta_1\\ \Theta_2\\ \Theta_3\end{bmatrix}^T \begin{bmatrix}X_1\\ X_2\\ X_3\end{bmatrix} + \begin{bmatrix}\Theta_4\\ \Theta_5\\ \Theta_6\end{bmatrix}^T \begin{bmatrix}{X_1}^2\\ {X_2}^2\\ {X_3}^2\end{bmatrix} + \begin{bmatrix}\Theta_7\\ \Theta_8\\ \Theta_9\end{bmatrix}^T \begin{bmatrix}X_1X_2\\ X_1X_3\\ X_2X_3\end{bmatrix} \qquad (1)$$

where, $\Theta_0, \Theta_1, \Theta_2, \Theta_3$ are linear effect coefficients; $\Theta_4, \Theta_5, \Theta_6$ are quadratic effect coefficients; $\Theta_7, \Theta_8, \Theta_9$ are interaction effect coefficients; e is the error. let's refer to the dependent variable - CO_2 concentration as C (in PPM) and the independent variables as, Ʈ (time in seconds), T (temperature in Celsius), and RH (relative humidity in %) respectively. The primary objective of this MVPR algorithm is to demonstrate the prediction equation shown in equation 1. To derive this prediction equation, it is necessary to identify all Θ_i coefficients. This regression's coefficients can be calculated by minimizing the cost function $J(\Theta)$ using the gradient descent approach. The cost function is the degree of error between prediction values and training dataset values. This approach employs the multivariate Mean Squared Error (MSE) Cost Function to compute the cost (4)(5)(6). By analyzing how accurately a predictor value represents the data, the regression model can adjust the coefficient values to get the optimal fit surface. In our case, the y, X_1, X_2, and X_3 are CO_2 concentration, relative humidity, temperature, and exposure time respectively. By substituting all these values in equ-1, the prediction equation of our case is formed, which follows as –

$$C = [\Theta_0 + e\,] + \begin{bmatrix}\Theta_1\\ \Theta_2\\ \Theta_3\end{bmatrix}^T \begin{bmatrix}RH\\ T\\ \text{Ʈ}\end{bmatrix} + \begin{bmatrix}\Theta_4\\ \Theta_5\\ \Theta_6\end{bmatrix}^T \begin{bmatrix}RH^2\\ T^2\\ \text{Ʈ}^2\end{bmatrix} + \begin{bmatrix}\Theta_7\\ \Theta_8\\ \Theta_9\end{bmatrix}^T \begin{bmatrix}RH.T\\ RH.\text{Ʈ}\\ T.\text{Ʈ}\end{bmatrix} \qquad (2)$$

Doing this optimization methodology manually on larger datasets is a complex and time-consuming process, therefore we employed computer programming to get the task done. Using the Python programming language, a code for this optimization problem of our Testo data was developed. This code employs NumPy, pandas, DateTime, seaborn, and matplotlib packages. Each coefficient in eq-2 has been characterized using this optimization process and substituted as shown in equation 3. The code of this task is provided in the supplementary materials section.

$$C = [-148475.938] + \begin{bmatrix}-4208.46\\ -599.844\\ 334.189\end{bmatrix}^T \begin{bmatrix}RH\\ T\\ \text{Ʈ}\end{bmatrix} + \begin{bmatrix}2.2\\ 318.2008\\ 0.00634\end{bmatrix}^T \begin{bmatrix}RH^2\\ T^2\\ \text{Ʈ}^2\end{bmatrix} + \begin{bmatrix}42.285\\ 2.587\\ 14.679\end{bmatrix}^T \qquad (3)$$

Equation-3 is the final regression-based model of prediction that we developed. This model can assist car occupants in estimating the amount of CO_2 that has accumulated in the cabin. This model can interpolate as well as extrapolate the dependent variables based on the inputs. The next section discusses the practical user interface features of this paradigm.

4. Model Representation

Implementing the developed model in the real world may be accomplished in several ways. A smart app, a website, or a graphical user interface (GUI) on the car's LED display might be appropriate methods for implementing our proposed model. Using these methods, automobile occupants may simply engage with our model and forecast the cabin's CO_2 concentrations. In response, we offered a GUI demonstration of this model's implementation to make it more comprehensible for readers. This demo GUI was created using the Python Tkinter package.

This GUI has a place for the user to type in values for temperature, relative humidity, and exposure time. Based on what the user puts in, our model can predict the concentration of CO_2 in the cabin with a "window opening message." For example, if people are in a car for 2 hours, they can put in the time, temperature, and relative humidity to figure out if they should open their windows to let in the fresh air. The demo GUI is set up such that it will show different messages at different levels of CO_2. These messages are based on the safe limits for CO_2 concentrations that different international organizations have set (7)(8). Based on the model's predicted value, the GUI will display the following notifications to the user.

- If CO_2 concentrations are between, 350-1000ppm: *"Predicted value is in the safe zone, you can close your windows"*
- If CO_2 concentrations are between, 1000-3000ppm: *"Open 25% of your cabin windows"*
- If CO_2 concentrations are between, 3000-5000ppm: *"Open 50% of your cabin windows"*
- If CO_2 concentrations are above 5000ppm: *"Open 75% of your cabin windows"*

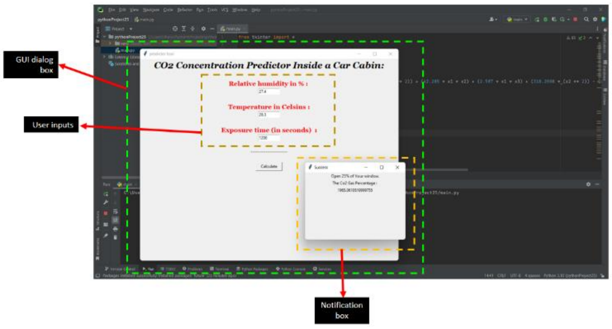

Figure 1: Demo Graphical user interface of our proposed model

Opening windows plays a significant part in improving the cabin's CO_2 levels since this allows outside air to enter the cabin and enhances ventilation. This ventilation is essential for optimizing indoor CO_2 levels and can lower the danger of airborne diseases. (1). The demo GUI is visually illustrated in Figure 1. The Tkinter code for this GUI is provided in the supplementary materials section. The GUI shown in Figure 1 is a demo version of the model representation, more work can be done in the form of mobile applications, websites, or car LED display screens to improve the user experience

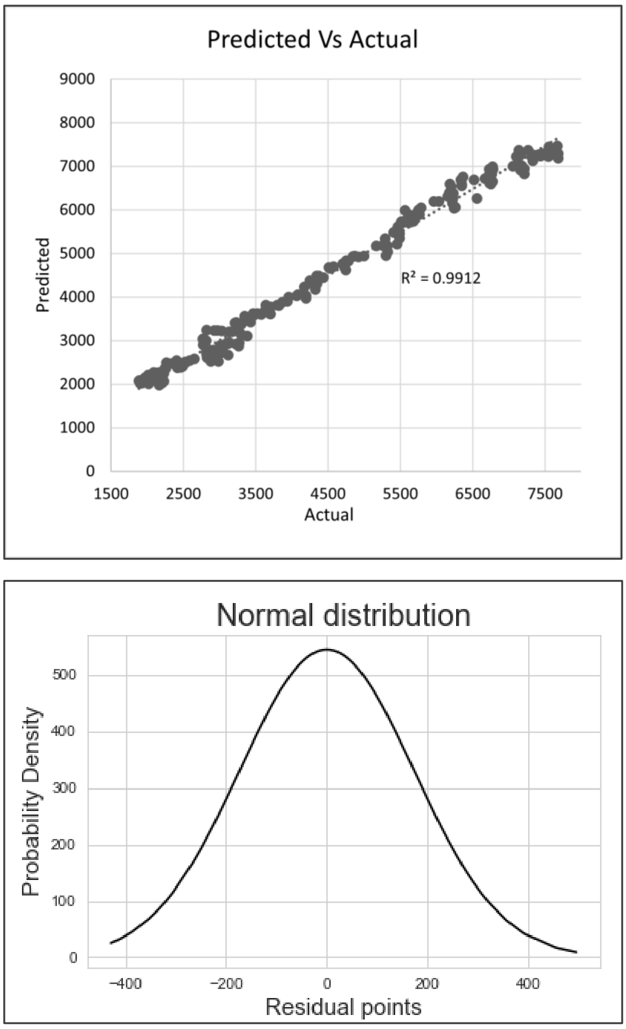

Figure 2: (a) correlation plot between Predicted and actual values (b) normal distribution of residuals

5. Model Validation

Using this deduced prediction equation (equ-3), the computed predicted values are compared to the actual values using Pearson's correlation theory (6). The correlation coefficient was calculated to be 99.12%, indicating that actual and predicted values are 99.12% more similar. This implies that the proposed model is accurate enough to predict CO2 concentrations within a car cabin. We further examined whether the proposed model corresponds to the residual normality assumption for the training dataset using Kolmogorov–Smirnov test in python. The KS test code is also provided in the supplementary materials. From, this test we concluded that the residuals are under a normal distribution and the distribution curve is shown in Figure 2 below.

6. Conclusions

Above-threshold amounts of accumulated CO2 gas are potentially hazardous for indoor occupants, particularly automobile drivers. The inhalation of this gas might impair the cognitive decision-making performance of automobile drivers; hence, it is essential to keep the cabin's concentration levels below the permitted limits. To do this, opening cabin windows can be an effective, simple, and economical method. However, opening cabin windows might always reduce thermal comfort. Consequently, opening windows for a little period can have a significant influence on optimizing the cabin's CO_2 concentration. In response, we developed a predicted model based on an MVPR machine learning algorithm and a demonstration GUI. This letter just presents preliminary results for this problem statement; in the future, further work and accurate monitoring may be done to more effectively address this problem statement.

7. Supplementary materials

The following link provides the supplementary materials for this study –
https://drive.google.com/drive/folders/1xJ2W0i8E1-pPAx5cMx9Ip_eK1LWmS60O?usp=sharing

Acknowledgment

This report's analysis forms part of our Btech final thesis project. Prof. Jay Dhariwal, department of design, Indian Institute of Technology, Delhi, monitored the training data mentioned in this study in Jodhpur, Rajasthan, India and we would like to thank him. Dr. Santhosh Kumar Mishra, the advisor of our Btech thesis titled "Statistics and experimental design for Production Engineers," is also deserving of our gratitude.

Ethical issue

Authors are aware of and comply with, best practices in publication ethics specifically about authorship (avoidance of guest authorship), dual submission, manipulation of figures, competing interests, and compliance with policies on research ethics. Authors adhere to publication requirements that the submitted work is original and has not been published elsewhere in any language.

Competing interests

The authors declare that no conflict of interest would prejudice the impartiality of this scientific work.

Authors' contribution

All authors of this study have a complete contribution to data collection, data analyses, and manuscript writing.

References

1. Klompas M, Baker MA, Rhee C. Airborne Transmission of SARS-CoV-2: Theoretical Considerations and Available Evidence. JAMA [Internet]. 2020;324(5):441–2. Available from: https://doi.org/10.1001/jama.2020.12458
2. Angelova RA, Markov DG, Simova I, Velichkova R, Stankov P. Accumulation of metabolic carbon dioxide (CO_2) in a vehicle cabin. IOP Conf Ser Mater Sci Eng. 2019;664(1).
3. Kajtár L, Herczeg L. Influence of carbon-dioxide concentration on human well-being and intensity of mental work. Időjárás. 2012 May 1;116:145–69.
4. FROST J. Curve Fitting Using Linear and Nonlinear Regression.[Online] Statistics by Jim [Internet]. 2018. Available from: https://statisticsbyjim.com/regression/curve-fitting-linear-nonlinear-regression/
5. Montgomery DC. Design and Analysis of Experiments Eighth Edition. Arizona State University. Vol. 2009, Copyright. 2013. 2001 p.
6. Costa V. Correlation and regression. In: Fundamentals of Statistical Hydrology [Internet]. The BMJ; 2016. p. 391–440. Available from: https://www.bmj.com/about-bmj/resources-readers/publications/statistics-square-one
7. Abdul-Wahab SA, Chin Fah En S, Elkamel A, Ahmadi L, Yetilmezsoy K. A review of standards and guidelines set by

international bodies for the parameters of indoor air quality. Atmos Pollut Res [Internet]. 2015;6(5):751–67. Available from: https://www.sciencedirect.com/science/article/pii/S1309104215301689

8. Jacobson TA, Kler JS, Hernke MT, Braun RK, Meyer KC, Funk WE. Direct human health risks of increased atmospheric carbon dioxide. Nat Sustain [Internet]. 2019;2(8):691–701. Available from: https://doi.org/10.1038/s41893-019-0323-1