Advanced Engineering Days

aed.mersin.edu.tr

Carbon monoxide forecasting with air quality parameters and fuzzy logic for Konya: A case study of Meram

Esra Aslı Çubukçu*[1], Vahdettin Demir [1], Mehmet Faik Sevimli [1]

[1]KTO Karatay University, Civil Engineering Department, Türkiye
cubukcuasli@gmail.com, vahdettin.demir@karatay.edu.tr, mehmet.faik.sevimli@karatay.edu.tr

Cite this study: Çubukçu, E. A., Demir, V., & Sevimli, M. F. (2022). Carbon monoxide forecasting with air quality parameters and fuzzy logic for Konya: A case study of Meram. 2nd Advanced Engineering Days, 65-68

Keywords
Air quality
Forecast
Fuzzy Logic
Konya
Carbon Monoxide

Abstract

The industrial revolution ushered in a period of fast technological advancement, which resulted in a surge in global consumerism. The phrase "air quality" has come to be associated with global climatic and environmental issues, particularly with industrialization around the world. Human activities have a higher impact on air pollution than natural events such as fires and volcanic eruptions. Air pollution has a variety of effects on natural and human life, including acid rain, respiratory disorders, and a decrease or rise in the number of living species in the environment. It is directly related to human health, particularly in cities that we might characterize as metropolises due to various industrial waste and gases. As a result, forecasting possible severe air pollution is critical for humans and the living ecosystem. The CO (μg/m^3) parameter of the Meram district in Konya province was evaluated using fuzzy logic methods in this study. PM_{10}, SO_2, NO_2, NO_X, and periodicity are input factors in the models that determine CO. CO is gas that is extremely damaging to human health. When the comparison criteria are analyzed, the FCM algorithm with five input models produces the best results.

1. Introduction

For many years, it has been well acknowledged that the quality of the air we breathe has a direct impact on our health. When the composition of the air changes in a way that disrupts human health or environmental balances, or when substances that shouldn't be in the air mix, it's termed air pollution. Nitrogen (N_2) makes up 78.084 percent of the air, whereas Oxygen (O_2) makes up 20.946 percent, Argon (Ar) makes up 0.934 percent, and carbon dioxide makes up 0.035 percent (CO_2). Neon (Ne), Methane (CH_4), Helium (He), Hydrogen (H_2), and Krypton (Kr) make up the remaining 0.001%. (Kr). With the increase in population, the growth of cities, and the development of industry, air pollution's effects are expanding at a faster rate and with a changing content. While air pollutants generated from a local source have local impacts, increased energy consumption, fossil fuel combustion, and the usage of motor vehicles all contribute to degradation of air quality in metropolitan areas. Regional transportation, acidification, increasing greenhouse gas emissions, and tropospheric ozone production all reflect the repercussions of air pollution that have already reached global proportions. While pollutants from traffic, transportation, industry, and heating (all of which are human) are the most common, the effects of meteorology, topography, and chemical transformation processes on air pollution and climate are now better understood. The impacts of air pollutants on the environment and human health are well understood to be dependent on time, space, duration of effect, concentration, and other factors. Because of heart and lung ailments, air pollution raises death rates. Furthermore, it has a harmful impact on children's lung development and raises the prevalence of chronic airway disorders such as asthma and chronic obstructive pulmonary disease (COPD) in polluted locations [1]. The Covid-19 illness epidemic overtook the globe in 2020. During the pandemic, it became even more apparent that the impact of air pollution on humans is rather significant. Covid-19 air pollution has been linked to an increased risk of getting the virus in numerous research investigations. Long-term air pollution also exposes millions of people to diseases of the respiratory and cardiovascular systems, such as diabetes, chronic diseases, and cancer. As a result, persons who are exposed to pollution are more vulnerable to viruses like Covid-19 [2]. Air pollution forecasting is vital to create individuals who are prepared and responsive to future extreme events[3, 4].

When some studies in the literature in recent years are examined; Unal et al. (2019) in their study, they used machine learning to forecast Ankara's air pollution. The findings were compared using RMSE, MAE, and R^2 [5]. Tunç and Toros (2020) in their study investigated into the impact of the Covid-19 epidemic on air pollution in Adana. While there was a minor drop in the concentration of environmental pollution both throughout the epidemic and compared to previous years, the results show that there was a slight increase in the concentration of environmental pollution during the pandemic. They discovered that the pandemic control measures had no substantial impact on the province of Adana's overall air quality [6]. Kara et al. (2020) evaluated NO, NO_2, NOx levels in Turkey by dividing them into Black Sea region, Marmara region, Aegean region, Mediterranean region, Central Anatolia region, and Eastern regions in their study done in 2020. Looking at the findings, it was discovered that the amount of nitrogen dioxide and its variations (NOx's) in Turkey has been decreasing over time [7]. Bayati et al. (2021) aimed to contribute to the improvement of air quality by monitoring the behavior of air pollutants over Van province through statistical analysis of the data; They discovered that the wind plays an important role in the horizontal transfer of contaminants in the atmosphere after analyzing the findings. They discovered that when the wind is calm, air pollution stays put, and precipitation aids in the collapse of contaminants in the atmosphere, making precipitation a cleaner of the atmosphere [8].

In this study the amount of CO (μg/m^3) in the air in Meram district was investigated using daily data from the Meram air quality station in Konya. In the modeling, three distinct Adaptive Neuro Fuzzy Inference System (ANFIS) were applied. There are used ANFIS-FCM, ANFIS-GP, and ANFIS-SC three methodologies. PM_{10}, SO_2, NO_2, NO_X, and periodicity (monthly) are the primary pollution constituents used as input parameters. CO is the output parameter. CO is a colorless, odorless gas that is produced when carbon in fuels is not entirely combusted, and it is extremely detrimental to human health. Internal combustion engines are the primary source (85-95 percent). CO emissions are mostly caused by industry, wood combustion, and forest fires. All measurements are in μg/m^3 unit.

2. Material and Method

2.1. Material

Air quality data were obtained from the air quality monitoring station in the Meram district of Konya, Turkey (https://sim.csb.gov.tr/STN/STN_Report/StationDataDownloadNew). 80% of the data was used in the training phase and 20% in the testing phase. The most important factor in choosing this period is the continuity of the data. The descriptive statistical information of the data used is given in Table 1.

Table 1. Statistical information of the data.

Statistic	PM_{10} (μg/m^3)	SO_2 (μg/m^3)	NO_2 (μg/m^3)	NO_X (μg/m^3)	CO (μg/m^3)
Average	28.46	12.47	38.75	60.55	908.20
Standard Eror	1.35	0.62	0.69	1.96	39.91
Median	20.46	7.62	37.07	50.42	601.30
Standard Deviation	24.76	11.40	12.64	35.93	729.53
Kurtosis	5.14	1.64	-0.22	4.97	5.05
Skewness	2.18	1.43	0.42	1.96	2.26
Maximum	3.29	1.26	12.22	15.00	205.80
Minimum	161.01	63.69	82.42	269.29	4201.67

2.2. Method

Fuzzy logic is built on the concept of subsets. An object is either a member of the set or it is not, according to the traditional approach. Fuzzy logic is a type of set theory that is based on classic logic. As a learning tool, clustering can be used to find data points in multivariate datasets. Clusters are separated into meaningful groupings. There are numerous different approaches to data clustering in diverse applications. When dealing with large datasets, however, the approach has significant limitations and may not perform as well as planned. On the other hand, other clustering algorithms have recently been created. One of the most widely used methods is fuzzy c-means clustering. For ANFIS- FCM; the data point is randomly initialized. Centers are selected using various clustering approaches. Minimizes errors by dividing the dataset. This algorithm continues iteratively until the convergence condition is met [9–11]. When employing the Grid Partition (GP) fuzzy inference system, the grid partitioning method is largely responsible for the model's learning. The Grid Partition approach separates the data set into rectangular sub-areas called grids based on the number and types of membership functions to be used. In the subspace, each input is partitioned into membership functions of the same shape. Based on input-output training, the system develops fuzzy rules that optimize data for quick learning and computation [12–14]. Sub Clustering approach (SC) is utilized when there is no clear understanding of the number of centers for data dispersion. This is one of the fuzzy clustering techniques. The number of possible cluster centers is completely determined by the number of data, not the dimensionality or distribution of the data. The data point with the closest neighbors is chosen as the cluster center. Other data points are positioned in the same way, with each point acting as a potential cluster center depending on its unique characteristics. This tactic is effective, especially because it is not dependent on other methods [13, 15].

3. Application

The comparative criteria were Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and Coefficient of Determination. In Equations 1-3, related equations are given.

$$MAE = \frac{1}{N}\sum_{i=1}^{N}\left|CO_e - CO_o\right| \qquad RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(CO_e - CO_o)^2} \qquad R^2 = \left(\frac{N*\left(\sum CO_o * CO_e\right) - \left(\sum CO_o\right)*\left(\sum CO_e\right)}{\sqrt{\left(N*\sum CO_o^2\right) - \left(\sum CO_e\right)^2 * \left(N*\sum CO_e^2\right) - \left(\sum CO_e\right)^2}}\right)^2$$

In the equations, “CO_e” and “CO_o” show the estimated and observed elevation values and “N” represents the number of data. Training and testing results are given in Table 2. The best results are colored red.

Table 2. Training and Testing Results

CRITERIA	METHODS	INPUTS				
		1 INPUT	2 INPUT	3 INPUT	4 INPUT	5 INPUT
MAE (Training)	FCM	183.713	125.212	121.650	113.853	114.124
	GP	383.773	407.148	374.075	398.713	63.205
	SC	179.128	119.973	126.523	79.622	120.082
RMSE (Training)	FCM	279.031	194.999	164.609	155.335	155.711
	GP	396.863	422.243	389.746	425.603	86.800
	SC	278.442	171.398	173.697	95.357	175.443
R^2 (Training)	FCM	0.877	0.940	0.957	0.962	0.962
	GP	0.892	0.214	0.356	0.004	0.768
	SC	0.877	0.954	0.952	0.854	0.951
MAE (Testing)	FCM	163.063	126.490	111.654	95.924	96.797
	GP	346.892	373.904	354.540	360.583	115.475
	SC	166.020	142.345	117.143	79.622	92.156
RMSE (Testing)	FCM	210.655	181.527	153.372	114.078	115.972
	GP	351.676	379.318	361.483	373.846	141.143
	SC	220.188	179.682	162.196	95.357	113.310
R^2 (Testing)	FCM	0.129	0.482	0.492	0.810	0.824
	GP	0.748	0.142	0.030	0.006	0.386
	SC	0.108	0.354	0.479	0.854	0.826

In Table 2, ANFIS-SC 4 (MAE=79.621, RMSE=95.356, R^2=0.853) input in the training phase and ANFIS-FCM 5 input (MAE=96.797, RMSE=115.972, R^2=0.824) result in the testing phase gived the best model.

4. Results

The amount of CO (μg/m^3) in the air was forecasted for Meram in this investigation. It was forecasted using daily data from the Meram air quality station in Konya's central district. The parameters in the model are primary air polluting gases.

In the modeling, three distinct Fuzzy logic algorithms used. These algorithms were FCM, GP, and SC. FCM 4 and 5 input models, as well as SC 5 input models, gives the best outcomes.

References

[1] URL: T. C. S. B. Sağlığı, “Hava Kirliliği ve Sağlık Etkileri,” *Halk Sağlığı Genel Müdürlüğü, Çevre Sağlığı Dairesi Başkanlığı.*, 2022. https://hsgm.saglik.gov.tr/tr/cevresagligi-ced/ced-birimi/hava-kirliliği-ve-sağlık-etkileri.html.

[2] URL : T. H. H. Platformu, “Kara Rapor 2021 Hava Kirliliği ve Sağlık Etkileri,” 2021. https://www.temizhavahakki.com/kararapor2021/

[3] Mills I. C., Atkinson R. W., Kang S., Walton H., and Anderson H. R., (2015). Quantitative systematic review of the associations between short-term exposure to nitrogen dioxide and mortality and hospital admissions.

BMJ Open, vol. 5, no. 5, doi: 10.1136/bmjopen-2014-006946.
[4] Saraçoğlu H., (2010). İzmir Limanına Gelen Gemilerin Oluşturduğu Egzoz Gazı Emisyonlarının İncelenmesi ve Çevresel Etkileri. Yüksek Lisans Tezi, İstanbul Teknik Üniversitesi, Fen Bilimleri Enstitüsü. İstanbul.
[5] Ünal Z. F., Dinç U., Özen C., and Toros H., (2019). Air Pollution Forecasting for Ankara with Machine Learning Method. J. Res. Atmos. Sci., 1(1), 42–48,.
[6] Tunç D. Ç. and Toros H., (2021). The Impact of COVID-19 on Air Quality in the city of Adana, Turkey. J. Res. Atmos. Sci., 2(2), 46–50. doi: 10.1080/15275922.2021.1892876.
[7] Kara Y., Karakaya T., Pirselimoğlu G., Dursun Ş., and Toros H., (2020) Overall Evaluation of NO x , NO, NO_2 Gasses in Turkey and Their Data Quality Control. Res. Artic. J. Res. Atmos. Sci., 2(1), 12–16.
[8] Al-bayati R. M., Bulut B., Adeeb H. Q., and Toros H., (2021). Air Pollution Data Analysis Over Van City , Turkey Van Hava Kirliliği Analizi. J. Res. Atmos. Sci., 3(1), 8–12.
[9] Jain, A. K. & Dubes R. C., (1988). Algorithms for clustering data. Prentice-Hall, Inc.
[10] Tsai, D. M. & Lin, C. C., (2011) Fuzzy C-means based clustering for linearly and nonlinearly separable data. Pattern Recognit., 44(8), 1750–1760. doi: 10.1016/j.patcog.2011.02.009.
[11] Bezdek, J. C., Ehrlich R., and Full W., (1984). FCM: The fuzzy c-means clustering algorithm. Comput. Geosci., 10(2–3), 191–203. doi: 10.1016/0098-3004(84)90020-7.
[12] Hekim M. and Orhan U., (2011). Bulanık c-means kümeleme yöntemine çıkarımlı yaklaşım. İtü Mühendislik Derg., 10–1, 356, 11–17.
[13] Fattahi H., (2016). Indirect estimation of deformation modulus of an in situ rock mass: an ANFIS model based on grid partitioning, fuzzy c-means clustering and subtractive clustering. Geosci. J., vol. 20, no. 5, pp. 681–690, doi: 10.1007/s12303-015-0065-7.
[14] Benmouiza K. & Cheknane, A., (2019). Clustered ANFIS network using fuzzy c-means, subtractive clustering, and grid partitioning for hourly solar radiation forecasting. Theor. Appl. Climatol., 137(1–2), 31–43. doi: 10.1007/s00704-018-2576-4.
[15] Priyono A., Ridwan M., Alias A. J., Rahmat R. A. O. K., Hassan, A., & Mohd. Ali, M. A., (2005). Generation of Fuzzy Rules with Subtractive Clustering. J. Teknol., 43(1), 143–153. doi: 10.11113/jt.v43.782.