



## Examining Air Pollution Continuity in Tehran Province using Markov Chain Model

A. Yousefi Kebriya, M. Nadi\*

Department of Water Engineering, Faculty of Agricultural Engineering, Sari Agricultural Sciences and Natural Resources University, Sari, Iran

### PAPER INFO

#### Paper history:

Received 25 August 2023

Accepted in revised form 05 October 2023

#### Keywords:

Air pollution

Air quality index

Markov chain

Transmission probability

### ABSTRACT

The Air Quality Index is a numerical tool used to quantify air pollution levels and classify pollution severity. It plays a vital role in ensuring healthcare system stability by understanding air pollution's dynamic behavior and shifts in pollution intensity. To analyze the probabilistic transition between pollution severity levels, a Markov Chain model was utilized. This study examined six air pollution states (Clean, Healthy, Unhealthy for Sensitive Groups, Unhealthy, Very Unhealthy, Hazardous) across 12 stations in Tehran's northern, southern, eastern, western, and central regions from 2018 to 2022. Results revealed that the western and southern areas exhibited the highest pollution levels, with over 44% and 34% of instances corresponding to unhealthy indices, respectively. In contrast, northern Tehran consistently maintained cleaner air quality. Probability transition matrices highlighted the greatest stability continuity in healthy air quality across all regions. Transitioning between clean/healthy air to very unhealthy/hazardous air was minimal, with hazardous air quality almost absent in most stations, except for the west and south (25% stability). The probability of continued unhealthy air quality in these areas reached 60%, indicating heightened pollution. The findings of transition probability matrices indicated that the western and southern regions had the highest likelihood of sustained pollution, while the northern region consistently maintained cleaner air. The probability of continuous clean air in the west was below 30%, while transitioning from very unhealthy/hazardous air to unhealthy air reached 50%. Conversely, the northern Tehran region exhibited over 40% stability for unhealthy air quality and over 50% for clean and healthy air.

doi: 10.5829/ijee.2024.15.02.10

### INTRODUCTION

Air pollution is a concerning issue, particularly for individuals residing in large industrialized cities, as it can have significant impacts on both human health and the environment. Rapid urbanization and industrialization worldwide have led to increasing severity of air quality. Efforts to meet the housing, consumption, healthcare, and energy needs of the growing urban population have resulted in numerous challenges, contributing to air pollution in modern cities. A comprehensive study conducted by the World Health Organization (WHO) revealed that nearly 90% of the global population is continuously exposed to air pollution, resulting in more than two million premature deaths annually [1]. Indoor Pollutant Sources Air pollution has an impact on the

body, including respiratory systems and heart disorders. Asthmatics, bronchiolitis, lung disease, cardiovascular problems, central nervous system malfunction, and skin conditions are the most common respiratory disorders, as are chronic obstructive pulmonary disease (COPD). The challenges posed by outdoor air pollution are public health risks such as cardiovascular disease, respiratory ailments, COPD, and world-class asymmetry [2]. Air quality in the residential areas adjacent to the industrial regions is of great concern due to the association with particulate matter and toxic gaseous which has adverse effects on human health [3]. Considering the influence of air on human life and its consumption overnight, it can be stated that the unfavorable quality of air in today's world is considered one of the most important environmental problems in many of the world's large cities [4]. The rapid

\*Corresponding Author Email: [mehdi.nadi@gmail.com](mailto:mehdi.nadi@gmail.com) (M. Nadi)

expansion of cities and urban populations, coupled with industrial growth, especially in proximity to large and densely populated cities, has not only escalated urban air pollution but has also triggered significant climatic and environmental changes, gaining substantial attention from researchers and planners [5].

Urbanization in recent decades has led to a significant increase in the number of cities with a population exceeding one million, rising from one city in 1976 to eight cities in 2017 in Iran alone. This population growth, coupled with inadequate urban planning, has brought about various environmental and human-related challenges, with air pollution being a prominent issue. Air pollution is just one of the many environmental problems arising from population growth and urban development. In Tehran, the capital of Iran, and several other Iranian cities, air pollution has become one of the most tangible environmental challenges [6]. Statistics show that during periods of increased air pollution in Tehran, the number of respiratory patients increases by up to 60%. The main factors associated with the rise in cardiovascular, respiratory, and pulmonary diseases are the increased levels of sulfur dioxide, particulate matter, and carbon monoxide [7]. Today, the various consequences of air pollution have necessitated continuous monitoring and control of air quality, making it an irrefutable concern in societies. Solving the complex issue of air pollution necessitates collective efforts and a deep understanding of pollution sources and events [8]. Pollution events, although not precisely predictable, can provide valuable insights when observed continuously. Probability laws play a crucial role in understanding many pollution events, contingent on knowledge of air quality indices and the probability of such events occurring [9]. In essence, the Air Quality Index, a vital tool for assessing and quantifying air quality, particularly for five major air pollutants—particulate matter, nitrogen dioxide, ground-level ozone, carbon monoxide, and sulfur dioxide—plays a pivotal role in comprehending the quality of air and the likelihood of pollution events.

The Air Quality Index (AQI) divides air quality into six categories: good, moderate, unhealthy for sensitive groups, unhealthy, very unhealthy, and hazardous. The likelihood of each category depends on the levels of pollutant concentrations [10]. According to probability laws, some random phenomena are more likely to occur when the probability of others is lower. In addition, among  $n$  possible states, only one state can occur, and the likelihood of each state is not necessarily superior to others. These processes are proven using Markov chains [11].

Utilizing Markov chains and understanding the continuity of pollution events and the possibility of creating various states can significantly aid in planning for air pollution management and reduction. Several studies have been conducted in this regard. Hoyos et al. [12] investigated the pollution of ozone and sulfur dioxide

over seven years (2006-2000) in Mexico City using a Markov chain-based model. The results showed that ozone had a steady distribution, while sulfur dioxide exhibited oscillatory behavior. Jarquin et al. [13] used Markov chain analysis to study the behavior of five pollutants ( $\text{CO}_2$ ,  $\text{NO}_x$ ,  $\text{SO}_2$ ,  $\text{O}_3$ ,  $\text{PM}_{10}$ ) for five possible states, ranging from good to very bad, in Mexico City. The results indicated that state changes over time exhibit nonlinear behavior, and even when pollutant concentrations increase progressively, the probability of worsening air quality is very low because it remains in an orderly state. Holmes and Hassini. [14] examined the random behavior of AQHI risk categories in Ontario (34 air monitoring stations) over five years from 2015 to 2019. They determined discrete-time Markov chains using three AQHI risk categories (low risk, moderate risk, high risk) as states for transition probabilities. Sarvi et al. [15] assessed the number of excess  $\text{PM}_{2.5}$  days in Tehran from October 2010 to December 2015. They used a hidden-state Poisson Markov model, considering various hidden states, to predict the number of occurrences. The results revealed that the number of days exceeding the limit (predicted average from 6.87 to 11.39 days) is relatively high for sensitive individuals based on  $\text{PM}_{2.5}$  air quality. Alyousifi et al. [16] applied a Markov Chain-based Fuzzy State (MCFS) model using triangular fuzzy numbers to analyze uncertainty in air pollution events and describe the transitional behavior of air pollution. Their results indicated that MCFS can effectively model air quality indices and may be a superior option for predicting air pollution. Zakaria et al. [17] conducted research on predicting future air quality using a Markov Chain model, introducing it as a model for long-term pollution distribution assessment. Suhaimi et al. [18] used the Multiple Imputation by Markov Chain Monte Carlo (MCMC) for air quality data calculation and found that the multiple imputation technique using the MCMC method provides a suitable fit and supplies missing data effectively.

Considering the significance of air quality in the urban areas of Tehran and the absence of prior statistical studies regarding the assessment of consecutive days with various air quality indices and the determination of the transition probabilities of pollution severity classes in Tehran and Iran, this research aims to analyze the frequency of pollution severity classes and investigate the likelihood of consecutive different air pollution severity levels in the city of Tehran.

Considering the utmost significance of urban air quality in Tehran and the substantial impact of air pollution on human health and the environment, coupled with the fact that comprehensive studies regarding the continuity of air pollution with various quality indices and the precise determination of the transition modes of different severity classes of pollution have not been conducted thus far for Tehran and other regions of Iran, this research aims to fill this crucial gap. It does so by

conducting an extensive analysis of the frequency of various air pollution severity categories and introducing a Markov chain model to analyze the persistence of these severity levels and the likelihood of consecutive days with varying air quality in Tehran. This research holds paramount importance in enhancing air quality, safeguarding human health, and protecting the environment.

**MATERIAL AND METHODS**

**Study area**

Tehran, the capital of Iran, is situated in the heart of the country with a population of approximately 8.5 million residents. This population can swell to over 12.5 million during the day, as people from surrounding cities commute to Tehran for work. There are over 17 million daily car trips in Tehran, with many vehicles which are outdated technology. Consequently, Tehran's air quality ranks among the most polluted in the world. The city's topography and climate exacerbate pollution issues. Tehran is located at a high altitude and is surrounded by the Alborz mountain range, which traps polluted air. Temperature inversions, especially in winter months, further hinder pollutant dispersion. Recent studies indicate that achieving clean air in Tehran is a complex challenge, as factors like rapid population growth, industrial development, urbanization, and increased fuel consumption exert significant pressure on air quality [19].

**Required data**

The ground-based air quality index (AQI) data was obtained from the Tehran Air Quality Control Company's website. There are 26 active monitoring stations in Tehran, measuring concentrations of air pollutants including ozone, particulate matter with diameters less than 10 and 2.5 micrometers, carbon monoxide, nitrogen dioxide, and sulfur dioxide. These monitoring stations determine the air quality index based on pollutant levels [20].

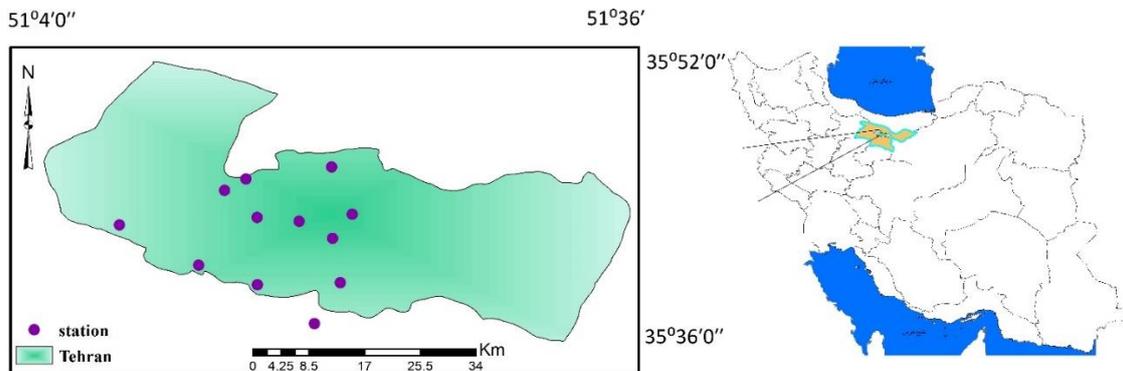
In this research, due to the need for daily time series data of the AQI for Markov analysis, data from 12 pollution monitoring stations in different areas of Tehran, including north, south, east, west, and central regions, were utilized. The data was collected over a 5-year period (2018-2022) from stations with minimal missing data. The coordinates of these monitoring stations are presented in Table 1, and the distribution of these stations is illustrated in Figure 1.

**Air quality index (AQI)**

The air quality index is a numerical scale used to report the daily air quality in terms of its impact on human health and the environment. The daily results of this index provide an estimate of air pollution levels that people rely on. An increase in the air pollution index signifies an increase in air pollution, posing a serious threat to human health. In most cases, the air quality index indicates how clean or polluted the air around us is. The primary focus of the air pollution index is on the health effects that may

**Table 1.** Geographical coordinates of ground pollution measuring stations

St	lat	long	z
Aghdasie District 14	35.79	51.48	1539
Ponak	35.76	51.33	1473
Pirouzi District 13	35.69	51.48	1208
Tarbiat modares District 6	35.72	51.38	1298
Setad bohran District 7	35.71	51.44	1271
Shad abad District 18	35.65	51.30	1139
Shahr rey District 20	35.57	51.462	1056
Shahrdari District 2	35.77	51.36	1558
Shahrdari District 19	35.63	51.38	1100
Shahrdari District 21	35.71	51.19	1234
Golbarg District 8	35.72	51.51	1289
Masoudie District 15	35.63	51.49	1179



**Figure 1.** Distribution of air pollution measurement stations in Tehran

be experienced days or hours after breathing polluted air. Calculations of the air pollution index are based on major air pollutants, including particulate matter, ground-level ozone, sulfur dioxide, nitrogen dioxide, and carbon monoxide. Particulate matter and ozone are the most hazardous to human health and the environment. For each of these categories of air pollutants, different countries have their own air quality index concerning other national air quality standards for public health protection, as summarized in Table 2 [21].

**Markov chain model**

Markov chain models can be used for simulating discrete-time series. The principles of this method are based on the temporal correlation between the data of a variable, meaning that the occurrence of a specific random variable, such as very unhealthy air quality, depends on the data of previous time steps or time increments. Each Markov chain model is determined by the number of states of the variable in question and its time order [22]. For example, in the analysis of air quality with a Markov chain, the number of model states can be set equal to the number of pollution levels. The model's order means that the state of a variable at the current time depends on how many time steps immediately preceding it, which can be 1 or 2 steps or more. The equation form of a first-order Markov chain is as follows:

$$Pr\{X_t|X_{t-1}, X_{t-2}, \dots, X_1\} = Pr\{X_t|X_{t-1}\} \tag{1}$$

The above equation indicates that the probability of an event such as X occurring at time t depends only on times t-1 (first order) or t-1 and t-2 (second order), not on the path that the time series has taken to reach the variable X at time t. In the following, some characteristics of first-order Markov chains, including transition probability matrices and the calculation of their steady-state probabilities, will be discussed [23].

Each element of the transition probability matrix in a first-order Markov chain model is denoted as P<sub>ij</sub>, representing the probability of transitioning from state i

to state j. This probability is determined using the following equation:

$$P_{ij} = \frac{n_{ij}}{\sum_{j=1}^S n_{ij}} \tag{2}$$

In this study, six states (equivalent to six air quality categories) were chosen for both i and j, as presented in Table 2. The form of the transition probability matrix for a first-order six-state Markov chain will be as follows:

$$P_{ij} = \begin{bmatrix} P_{11} & P_{12} & P_{13} & P_{14} & P_{15} & P_{16} \\ P_{21} & P_{22} & P_{23} & P_{24} & P_{25} & P_{26} \\ P_{31} & P_{32} & P_{33} & P_{34} & P_{35} & P_{36} \\ P_{41} & P_{42} & P_{43} & P_{44} & P_{45} & P_{46} \\ P_{51} & P_{52} & P_{53} & P_{54} & P_{55} & P_{56} \\ P_{61} & P_{62} & P_{63} & P_{64} & P_{65} & P_{66} \end{bmatrix} \tag{3}$$

In the above matrix, for example, the probability that after clean air, unhealthy air occurs. In this matrix, the sum of probabilities in each row equals 1, meaning:

$$\sum_{j=1}^S P_{ij} = 1 \tag{4}$$

To determine the time dependence (first and second orders) or the lack of time dependence (zero order) in discrete time series of AQI, it is necessary to use an appropriate criterion. For this purpose, the Bayesian Information Criterion (BIC) is commonly used [24].

$$BIC(m) = -2L_m + s^m(\ln n) \tag{5}$$

In which, S represents the number of states, m is the order of the model, and the parameter is determined from the following relationships [25]:

$$L_0 = \sum_{j=0}^{S-1} n_j \ln(P_j) \tag{6}$$

$$L_2 = \sum_{h=0}^{S-1} \sum_{i=0}^{S-1} \sum_{j=0}^{S-1} n_{hij} \ln(P_{hij}) \tag{7}$$

In which, the parameter represents the likelihood parameter of the model of order m. The model with a lower BIC value is introduced as the optimal model. If the BIC values are equal for several different models, the model with a lower order is selected [26].

**Table 2.** Air quality index range and air pollution class

Description of air quality	Range of AQI	Levels of concern	Class number
Air quality is satisfactory, and air pollution poses little or no risk	0-50	Good	1
Air quality is acceptable. However, there may be a risk for some people, particularly those who are unusually sensitive to air pollution.	51-100	Moderate	2
Members of sensitive groups may experience health effects. The general public is less likely to be affected.	101-150	Unhealthy for Sensitive Groups	3
Some members of the general public may experience health effects; members of sensitive groups may experience more serious health effects.	151-200	Unhealthy	4
Health alert: The risk of health effects is increased for everyone.	201-300	Very Unhealthy	5
Health warning of emergency conditions: everyone is more likely to be affected.	301 and higher	Hazardous	6

**RESULTS AND DISCUSSION**

**Selection the appropriate order of the Markov chain**

In this research, daily data for five years (2018-2022) from 12 air quality monitoring stations in five regions: East, West, North, South, and Central Tehran were used to select the appropriate order for the Markov chain model. The investigation of orders 0, 1, and 2 using the BIC index in Table 3 showed that the first-order model had the lowest BIC value and therefore, the best Markov model for fitting all the pollution monitoring stations in Tehran is a first-order Markov chain model with six states.

**Frequency analysis of air quality index**

Based on the analysis of the air quality frequency index in Tehran, as provided in Table 4, the western regions are the most heavily polluted areas, displaying the highest percentages of both healthy and very unhealthy air quality categories. For instance, Shad Abad and Shahr-dari 21 Municipality stations show healthy and clean air quality index percentages of 55.9% and 51.4%, respectively. Furthermore, more than 44% of days are associated with unhealthy air pollution conditions in these areas.

The second most polluted area in Tehran is situated in the southern region, with the unhealthy air quality index percentages at Shahr Rey and Shahr-dari 19 stations being 34.91% and 31.34%, respectively.

In terms of the pollution index percentage, the central region of Tehran, particularly the Tarbiat Modares station, surpasses even the southern region, with the pollution index percentage exceeding 40%. However, at the Setad Bohran station, this figure is 31.5%.

The eastern areas of Tehran exhibit better air quality compared to the western and southern areas. The combined percentage of unhealthy air quality categories at Golbarg and Masoudie stations is 22.2% and 16.81%, respectively. Conversely, the Pirouzi region is more polluted than the southern regions, with the pollution index percentage at the Pirouzi station exceeding 39%, indicating higher pollution levels compared to other eastern regions of Tehran. Clean and healthy air conditions, as indicated by the pollution index frequency, are prevalent in the northern region of Tehran. Here, less than 29% of days in the year experience polluted conditions, and on more than 70% of days in the year, the air quality in this area is considered clean and healthy.

**Probabilistic analysis of daily transitions with different AQI**

Given that in the majority of AQI index stations and time intervals, the first-order Markov model provides a better fit for the Air Quality Index (AQI), selected the first-order Markov chain model with six states for analyzing air quality classes. The transition probability matrix for this model is a square 6\*6 matrix, with each element

**Table 3.** BIC index values of Markov chain models of different orders in Tehran air quality monitoring stations

Order	Aghdasie	Ponak	Pirouzi	Tarbiat modares	Setad bohran	Shad abad	Shahr	Shahr-dari2	Shahr-dari19	Shahr-dari21	Golbarg	Masoudie
0	4207	3476	3788	4034	4032	3822	3411	4016	4343	3816	3763	3200
1	3089	2831	3041	3280	3271	3143	2892	3054	3204	3032	2922	2475
2	3282	2923	3198	3431	3384	3282	3026	3190	3247	3101	3035	2687

**Table 4.** Frequency (%) analysis of air quality of studied stations in Tehran city

Station	0-50	51-100	101-150	151-200	201-300	301<	Geographical direction
Tarbiat modares	3.4	56.4	26.0	13.2	0.9	0.1	Central
Setad bohran	8.4	60.2	21.8	8.9	0.7	0.1	Central
Golbarg	14.3	63.6	17.6	4.1	0.4	0.1	East
Masoudie	12.9	70.0	14.9	1.8	0.1	0.01	East
Pirouzi	2.7	57.9	26.8	11.6	0.1	0.16	East
Aghdasie	15.3	56.0	21.3	6.9	0.4	0.06	North
Ponak	13.5	65.6	18.7	2.2	0.1	0.05	North
Shahr-dari2	12.4	60.4	19.4	7.3	0.3	0.11	North
Shahr rey	2.6	62.6	28.5	6.1	0.1	0.11	South
Shahr-dari19	11.4	54.1	24.2	9.6	0.4	0.22	South
Shad abad	1.9	54.0	32.1	11.6	0.1	0.3	West
Shahr-dari21	1.3	50.1	34.3	11.7	1.38	1.22	West

representing the probability of transitioning from one specific air quality class to another.

To gain a deeper understanding of air quality in various regions of Tehran, including the east, west, north, south, and central areas, we conducted probabilistic analyses at multiple monitoring stations.

**Central Region of Tehran:** The results presented in Table 5 reveal that the central regions of Tehran exhibit the highest level of stability and consistency for AQI class 2, which signifies healthy air. The probability of transitioning from one day with healthy air to the next day with healthy air stands at 75%. Additionally, there is a probability exceeding 40% of the air quality remaining unhealthy for sensitive groups (AQI class 3).

Table 5 results indicate that there is no possibility of transitioning from clean air to unhealthy, very unhealthy, or hazardous air, and vice versa. Nevertheless, there is a likelihood of transitioning from days with a high pollution index to unhealthy air and air quality indices for sensitive groups. Furthermore, the stability of indices denoting unhealthy air for sensitive groups and unhealthy air for all groups surpasses the 50% mark. Conversely, there is no stability observed in the very hazardous air quality index, while the stability of the hazardous air quality index stands at 50%.

Table 5 further illustrates that in instances of unhealthy, very unhealthy, and hazardous air pollution conditions, there exists a probability that the air quality will persist as unhealthy or transition to air quality indices designed for sensitive groups on the following day.

**Eastern Region of Tehran:** Considering that the main diagonal of the transition probability matrix represents the continuity and stability of each air pollution class, our examination of the main diagonal elements in the probability matrices for air pollution class transitions at stations in the eastern part of Tehran (as presented in Table 6) reveals noteworthy findings.

Firstly, it's evident that the Golbarg and Masoudie stations exhibit high probabilities of maintaining clean and healthy air, with probabilities exceeding 55% and 75%, respectively. The Pirouzi station also displays substantial continuity in clean and particularly healthy air, with probabilities of 29% and 77%. These observations

underscore the remarkable persistence of clean and, notably, healthy air quality in these locations.

Interestingly, Golbarg station indicates a probability of more than 45% for days characterized by an unhealthy air quality index. In contrast, the Masoudie station shows no instances of very unhealthy or hazardous air pollution persisting. However, it's worth noting that there is a higher probability of air quality transitioning from these categories to unhealthy conditions and those specifically affecting sensitive groups. Consequently, the likelihood of a day with severe air pollution transitioning to a day with clean air is low, with a probability of less than 50%. Conversely, the reverse scenario, where healthy conditions lead to unhealthy conditions, has not been observed.

Furthermore, it appears that the Masoudie station consistently experiences healthier weather conditions compared to Golbarg and Pirouzi stations.

These insights provide valuable information for understanding the air quality dynamics in the eastern region of Tehran, contributing to informed decision-making in environmental management and public health. **Northern Tehran Region:** The outcomes derived from the probability matrices detailing air pollution level transitions in the northern regions of Tehran are presented in Table 7. A comprehensive analysis of these matrices reveals several significant patterns and trends.

Firstly, it's apparent that the probability of air quality remaining unhealthy in northern Tehran exceeds 40%, while clean and healthy air quality classes boast a higher probability, surpassing the 50% mark. This highlights the region's capacity for sustaining clean and healthy air quality conditions.

Notably, the probability of transitioning from days characterized by hazardous pollution levels to unhealthy air quality indices for sensitive groups is notably higher compared to other transitions. Conversely, the possibility of transitioning to a day with clean and healthy air quality following a hazardous pollution day is non-existent. In essence, once hazardous air pollution conditions are experienced on a particular day, there is no possibility of their recurrence on the subsequent day, and certainly, clean and healthy air conditions will not prevail.

**Table 5.** Transition probability matrices of central Tehran (A: Tarbiat modares, B: Setad bohran)

A							B						
Index	1	2	3	4	5	6	Index	1	2	3	4	5	6
1	<b>0.30</b>	0.63	0.06	0.00	0.00	0.00	1	<b>0.41</b>	0.58	0.02	0.00	0.00	0.00
2	0.04	<b>0.75</b>	0.19	0.03	0.00	0.00	2	0.08	<b>0.75</b>	0.16	0.01	0.00	0.00
3	0.01	0.39	<b>0.44</b>	0.16	0.01	0.00	3	0.02	0.42	<b>0.39</b>	0.16	0.02	0.00
4	0.00	0.15	0.28	<b>0.55</b>	0.02	0.00	4	0.01	0.10	0.36	<b>0.50</b>	0.02	0.00
5	0.00	0.00	0.13	0.38	<b>0.50</b>	0.00	5	0.00	0.17	0.33	0.42	<b>0.00</b>	0.08
6	0.00	0.00	0.90	0.10	0.00	<b>0.00</b>	6	0.00	0.00	0.50	0.50	0.00	<b>0.00</b>

**Table 6.** Transition probability matrices of east Tehran (A: Golbarg, B: Masodie, C: Pirouzi)

A							B							C						
Index	1	2	3	4	5	6	Index	1	2	3	4	5	6	Index	1	2	3	4	5	6
1	<b>0.59</b>	0.40	0.01	0.00	0.00	0.00	1	<b>0.59</b>	0.39	0.01	0.00	0.00	0.00	1	<b>0.29</b>	0.58	0.10	0.02	0.00	0.00
2	0.09	<b>0.78</b>	0.12	0.01	0.00	0.00	2	0.07	<b>0.82</b>	0.10	0.01	0.00	0.00	2	0.03	<b>0.77</b>	0.19	0.02	0.00	0.00
3	0.02	0.42	<b>0.47</b>	0.08	0.00	0.00	3	0.01	0.44	<b>0.49</b>	0.06	0.00	0.00	3	0.01	0.40	<b>0.45</b>	0.14	0.00	0.00
4	0.00	0.16	0.35	<b>0.45</b>	0.04	0.00	4	0.00	0.38	0.41	<b>0.19</b>	0.03	0.00	4	0.00	0.10	0.29	<b>0.59</b>	0.00	0.01
5	0.00	0.14	0.29	0.29	<b>0.29</b>	0.00	5	0.00	0.50	0.00	0.50	<b>0.00</b>	0.00	5	0.00	0.50	0.00	0.00	<b>0.00</b>	0.50
6	0.00	0.00	0.00	0.90	0.10	<b>0.00</b>	6	0.00	0.00	0.00	0.90	0.10	<b>0.00</b>	6	0.00	0.00	0.90	0.10	0.00	<b>0.00</b>

The transition probability matrices for stations in northern Tehran effectively illustrate that as air pollution intensity escalates, its stability diminishes. When hazardous air pollution conditions occur, there is a 50% probability of these conditions persisting as unhealthy air quality or unhealthy air quality for sensitive groups on the following day.

These findings provide valuable insights into the dynamics of air quality in the northern region of Tehran, emphasizing the challenges associated with addressing and mitigating air pollution, especially when pollution levels reach hazardous thresholds. This knowledge contributes to the development of targeted environmental and public health strategies.

**Southern Tehran Region:** In the southern Tehran region, akin to other areas, the probability of transitioning from the first four levels of air pollution, which include clean, healthy, unhealthy for sensitive groups, and unhealthy, to very unhealthy and hazardous conditions is nearly negligible. However, a notable distinction emerges in the Shahr-dari 19 station, where the probabilities of transitioning to dangerous, unhealthy, and unhealthy conditions for sensitive groups are 25%, 60%, and 53%, respectively. These probabilities exceed those observed at central, eastern, and northern stations. Moreover, there is approximately a 50% likelihood of transitioning from dangerous and very unhealthy air pollution conditions to unhealthy conditions. This observation underscores the

persistence of air pollution over consecutive days in the southern Tehran region. Comparing the air pollution transition probability matrices in the southern Tehran region, as presented in Table 8, reveals that in the Shahr Rey area, although the stability of clean air is lower compared to the Shahr-dari 19 station.

This analysis provides valuable insights into the unique dynamics of air quality in the southern region of Tehran, emphasizing the challenges posed by the persistence of air pollution, especially in comparison to other areas of the city.

**Western Tehran Region:** The probability matrices depicting air pollution transmissions in the western region of Tehran are detailed in Table 9, revealing distinct patterns that distinguish this area from others in the city.

Notably, the persistence of air pollution in the western region of Tehran, particularly unhealthy air, surpasses that of other areas in the city. Moreover, there exists a noteworthy 50% probability that after experiencing very unhealthy and hazardous air pollution conditions on a given day, unhealthy air conditions will prevail on the subsequent day. This observation underscores a concerning trend of consecutive days with air pollution in the western area of Tehran.

In contrast, the probability of sustained clean air in this region is considerably low, measuring below 30%. Additionally, there is no possibility of transitioning from

**Table 7.** Transition probability matrices of north Tehran (A: Aghdasie, B: Ponak, C: Shahr-dari 2)

A							B							C						
Index	1	2	3	4	5	6	Index	1	2	3	4	5	6	Index	1	2	3	4	5	6
1	<b>0.59</b>	0.40	0.01	0.00	0.00	0.00	1	<b>0.53</b>	0.47	0.00	0.00	0.00	0.00	1	<b>0.59</b>	0.40	0.01	0.00	0.00	0.00
2	0.11	<b>0.75</b>	0.13	0.01	0.00	0.00	2	0.09	<b>0.78</b>	0.12	0.00	0.00	0.00	2	0.08	<b>0.77</b>	0.14	0.01	0.00	0.00
3	0.01	0.32	<b>0.51</b>	0.16	0.01	0.00	3	0.02	0.41	<b>0.51</b>	0.06	0.00	0.00	3	0.02	0.39	<b>0.45</b>	0.14	0.00	0.00
4	0.00	0.13	0.41	<b>0.43</b>	0.03	0.00	4	0.00	0.15	0.49	<b>0.37</b>	0.00	0.00	4	0.00	0.16	0.30	<b>0.52</b>	0.01	0.01
5	0.00	0.25	0.25	0.38	<b>0.13</b>	0.00	5	0.00	0.00	0.00	0.90	<b>0.10</b>	0.00	5	0.00	0.20	0.00	0.40	<b>0.40</b>	0.00
6	0.00	0.00	0.90	0.10	0.00	<b>0.00</b>	6	0.00	0.00	0.90	0.10	0.00	<b>0.00</b>	6	0.00	0.00	0.90	0.10	0.00	<b>0.00</b>

**Table 8.** Transition probability matrices of south Tehran (A: Shahr Rey, B: Shahr-dari 19)

A							B						
Index	1	2	3	4	5	6	Index	1	2	3	4	5	6
1	<b>0.29</b>	0.67	0.04	0.00	0.00	0.00	1	<b>0.60</b>	0.38	0.02	0.00	0.00	0.00
2	0.03	<b>0.77</b>	0.19	0.01	0.00	0.00	2	0.08	<b>0.75</b>	0.16	0.01	0.00	0.00
3	0.01	0.42	<b>0.48</b>	0.09	0.00	0.00	3	0.01	0.33	<b>0.53</b>	0.12	0.01	0.00
4	0.00	0.12	0.39	<b>0.48</b>	0.01	0.01	4	0.01	0.13	0.24	<b>0.60</b>	0.02	0.01
5	0.00	0.00	0.90	0.10	<b>0.00</b>	0.00	5	0.00	0.00	0.38	0.50	<b>0.00</b>	0.13
6	0.00	0.00	0.50	0.50	0.00	<b>0.00</b>	6	0.00	0.00	0.25	0.50	0.00	<b>0.25</b>

unhealthy conditions to clean air conditions, both for the general population and sensitive groups. The probability of transitioning from AQI class 3 and 4 to healthy air conditions stands at less than 40% and 10%, respectively.

These findings shed light on the distinctive air quality dynamics in the western region of Tehran, emphasizing the challenges associated with mitigating air pollution and the need for targeted interventions to improve air quality and public health in this area.

Comparing the results with previous research, Rahimi et al. [27]. noted that CO pollutant levels were highest in Fatemi, Bazar, and Aghdasie stations throughout most months of the year, with Fatemi station showing the highest probability of CO pollution occurrence. Conversely, Shahr Rey station exhibited the lowest likelihood of CO pollutant persistence.

Building upon these insights, Holmes and Hassini [14] employed the Air Quality Health Index (AQHI) in eastern Canada, highlighting that urban areas spent more time in high-risk categories and tended to remain in these categories for extended durations before transitioning.

The examination of various orders of the Markov chain model revealed that pollution intensity over two consecutive days was not independent, suggesting a connection between air quality on consecutive days. The first order of the Markov chain was selected, implying that air quality on one day is influenced by the preceding day, while the air quality of two previous days does not

significantly impact the pollution intensity of the current day.

The general examination of transmission probability matrices demonstrated that, overall, as pollution intensity increased, the stability of air quality decreased. However, in most stations, no stability probability was observed for the dangerous air quality category. Notably, the probability of two consecutive days of healthy air quality was higher than that of consecutive days with clean air quality.

To further understand the AQI patterns, transmission probability matrices for air pollution monitoring stations were estimated. The results revealed that unhealthy air quality categories at each monitoring station exhibited stability, with gradual and low probability transitions between unhealthy and healthy categories. Sudden and significant transitions between very unhealthy or dangerous air quality to clean and healthy conditions, and vice versa, were not anticipated.

In a similar study, Zakaria et al. [17] employed the Markov chain to predict air quality in Sarawak, Malaysia. Their findings indicated a higher continuity of good and dangerous air quality conditions compared to the other four conditions. Additionally, the probability of good air quality conditions in the long term exceeded that of average and unhealthy states. Very unhealthy and dangerous conditions showed the lowest probability of continuation.

**Table 9.** Transition probability matrices of west Tehran (A: Shad abad, B: Shahr-dari 21)

A							B						
Index	1	2	3	4	5	6	Index	1	2	3	4	5	6
1	<b>0.26</b>	0.60	0.11	0.03	0.00	0.00	1	<b>0.25</b>	0.58	0.13	0.04	0.00	0.00
2	0.02	<b>0.73</b>	0.22	0.02	0.00	0.00	2	0.02	<b>0.71</b>	0.25	0.02	0.00	0.00
3	0.01	0.38	<b>0.51</b>	0.10	0.00	0.00	3	0.00	0.36	<b>0.53</b>	0.10	0.00	0.00
4	0.00	0.10	0.29	<b>0.60</b>	0.00	0.01	4	0.00	0.08	0.29	<b>0.61</b>	0.00	0.01
5	0.00	0.00	0.50	0.00	<b>0.00</b>	0.50	5	0.00	0.00	0.50	0.50	<b>0.00</b>	0.00
6	0.00	0.00	0.50	0.50	0.00	<b>0.00</b>	6	0.00	0.00	0.50	0.25	0.00	<b>0.25</b>

## CONCLUSION

In this research study, an extensive analysis of the Air Quality Index (AQI) across six categories, including clean, healthy, unhealthy for sensitive groups, unhealthy, very unhealthy, and dangerous levels, was conducted. The analysis focused on the probability of consecutive-day transitions for 12 monitoring stations situated across different regions of Tehran, encompassing the north, south, east, west, and central areas. The study spanned a five-year period from 2018 to 2022.

The findings revealed that the northern region of Tehran consistently exhibited the best air quality, while the western and southern regions consistently ranked as having the poorest air quality.

In conclusion, the probability of consecutive days with clean air quality in western Tehran was found to be less than 30%, whereas in northern Tehran, there was a probability of more than 60% for consecutive days with clean and healthy air quality. Importantly, it was observed that dangerous air pollution conditions, which are harmful to all living beings and associated with severe consequences, did not persist for more than two consecutive days in the north, east, and central regions of Tehran. However, there was a possibility of such conditions persisting in the southern and western regions, particularly in Shahrdari 19 and 21 areas.

These findings provide valuable insights into the dynamics of air quality in Tehran, underlining the importance of region-specific air quality management strategies and interventions.

## CONFLICT OF INTEREST

The authors declare that they have no conflict of interest.

## REFERENCES

- Manisalidis, I., Stavropoulou, E., Stavropoulos, A., and Bezirtzoglou, E., 2020. Environmental and Health Impacts of Air Pollution: A Review. *Frontiers in Public Health*, 8(14), pp.1–13. Doi: 10.3389/fpubh.2020.00014
- Singh, A., and Singh, K.K., 2022. An Overview of the Environmental and Health Consequences of Air Pollution. *Iranica Journal of Energy and Environment*, 13(3), pp.231–237. Doi: 10.5829/IJEE.2022.13.03.03
- Mohamed, R., Al-Gheethi, A.A., Fahira, M.A.B., Fahimah, H., Yahaya, N.Z., and Amir, H.K., 2017. Determination of Particulate Matter and Toxic Gaseous in Ambient Air Adjacent to Industrial Area. *Iranica Journal of Energy and Environment*, 8(2), pp.170–180. Doi: 10.5829/IJEE.2017.08.02.11
- Poormolaie, Z., Mohammadi, M., Ghafouri, M., and Khayyami, E., 2022. Determination of Air Quality Index and Its Relationship With Meteorological Parameters in City of Mashhad. *Iranian Journal of Energy and Environment*, 13(3), pp.273–283. Doi: 10.5829/IJEE.2022.13.03.07
- Glencross, D.A., Ho, T.-R., Camiña, N., Hawrylowicz, C.M., and Pfeffer, P.E., 2020. Air pollution and its effects on the immune system. *Free Radical Biology and Medicine*, 151, pp.56–68. Doi: 10.1016/j.freeradbiomed.2020.01.179
- Sharifi, A., Khavarian-Garmsir, A.R., Allam, Z., and Asadzadeh, A., 2023. Progress and prospects in planning: A bibliometric review of literature in Urban Studies and Regional and Urban Planning, 1956–2022. *Progress in Planning*, 173, pp.100740. Doi: 10.1016/j.progress.2023.100740
- Pantelic, J., Nazarian, N., Miller, C., Meggers, F., Lee, J.K.W., and Licina, D., 2022. Transformational IoT sensing for air pollution and thermal exposures. *Frontiers in Built Environment*, 8. Doi: 10.3389/fbuil.2022.971523
- Pouyanfar, N., Harofte, S.Z., Soltani, M., Siavashy, S., Asadian, E., Ghorbani-Bidkorbeh, F., Keçili, R., and Hussain, C.M., 2022. Artificial intelligence-based microfluidic platforms for the sensitive detection of environmental pollutants: Recent advances and prospects. *Trends in Environmental Analytical Chemistry*, 34, pp.e00160. Doi: 10.1016/j.teac.2022.e00160
- Maleki, H., Sorooshian, A., Goudarzi, G., Baboli, Z., Tahmasebi Birgani, Y., and Rahmati, M., 2019. Air pollution prediction by using an artificial neural network model. *Clean Technologies and Environmental Policy*, 21(6), pp.1341–1352. Doi: 10.1007/s10098-019-01709-w
- Weissert, L.F., Salmond, J.A., Miskell, G., Alavi-Shoshtari, M., and Williams, D.E., 2018. Development of a microscale land use regression model for predicting NO<sub>2</sub> concentrations at a heavy trafficked suburban area in Auckland, NZ. *Science of The Total Environment*, 619–620, pp.112–119. Doi: 10.1016/j.scitotenv.2017.11.028
- Greenland, S., Senn, S.J., Rothman, K.J., Carlin, J.B., Poole, C., Goodman, S.N., and Altman, D.G., 2016. Statistical tests, P values, confidence intervals, and power: a guide to misinterpretations. *European Journal of Epidemiology*, 31(4), pp.337–350. Doi: 10.1007/s10654-016-0149-3
- Hoyos, L., Lara, P., Ortiz, E., López, R., and González, J., 2009. Evaluation of air pollution control policies in Mexico City using finite Markov chain observation model. *Revista de Matemática: Teoría y Aplicaciones*, 16(2), pp.255–266. Doi: 10.15517/rmta.v16i2.305
- Jarquín, B., Aguilar Fernandez, M., Calderon Vidales, I. and Jiovanna Zamudio, J., 2021. Markov chains model applied to the analysis of air quality, International Journal of Latest Research in Science and Technology, 10(1), pp. 1-5. Retrieved from: <https://www.mnkjournals.com/journal/ijlrst/index.php>
- Holmes, J., and Hassini, S., 2021. Discrete-Time Markov Chain Modelling of the Ontario Air Quality Health Index. *Water, Air, & Soil Pollution*, 232(4), pp.158. Doi: 10.1007/s11270-021-05096-1
- Sarvi, F., Nadali, A., Khodadost, M., Kharghani Moghaddam, M., and Sadeghifar, M., 2017. Application of Poisson Hidden Markov Model to Predict Number of PM<sub>2.5</sub> Exceedance Days in Tehran During 2016-2017. *Avicenna Journal of Environmental Health Engineering*, 4(1), pp.58031–58031. Doi: 10.5812/ajehe.58031
- Alyousifi, Y., Kırıl, E., Uzun, B., and Ibrahim, K., 2021. New Application of Fuzzy Markov Chain Modeling for Air Pollution Index Estimation. *Water, Air, & Soil Pollution*, 232(276), pp.123–135. Doi: 10.1007/s11270-021-05172-6
- Zakaria, N.N., Othman, M., Sokkalingam, R., Daud, H., Abdullah, L., and Abdul Kadir, E., 2019. Markov Chain Model Development for Forecasting Air Pollution Index of Miri, Sarawak. *Sustainability*, 11(51), pp.65–80. Doi: 10.3390/su11195190
- Suhaimi, N., Ghazali, N.A., Nasir, M.Y., Mokhtar, M.I.Z., and Ramli, N.A., 2017. Markov Chain Monte Carlo method for handling missing data in air quality datasets. *Malaysian Journal of Analytical Sciences*, 21(3), pp.552–559. Doi: org/10.17576/mjas-

2017-2103-05

19. Shahbazi, H., Hosseini, V., and Rashidi, Y., 2014. Investigating The Impact of initial and Boundary Concentrations on Air Quality Modeling Accuracy over Tehran. *Environmental Sciences*, 12(3), pp.93–100. [In Persian]. Retrieved from: [https://envs.sbu.ac.ir/article\\_97440.html](https://envs.sbu.ac.ir/article_97440.html)
20. Ebrahimi, M., and Qaderi, F., 2021. Determination of the most effective control methods of SO<sub>2</sub> Pollution in Tehran based on adaptive neuro-fuzzy inference system. *Chemosphere*, 263(18), pp.120–135. Doi: 10.1016/j.chemosphere.2020.128002
21. Zhong, S., Yu, Z., and Zhu, W., 2019. Study of the Effects of Air Pollutants on Human Health Based on Baidu Indices of Disease Symptoms and Air Quality Monitoring Data in Beijing, China. *International Journal of Environmental Research and Public Health*, 16(6), pp.1–19. Doi: 10.3390/ijerph16061014
22. Wilks, D.S., 2011. Statistical methods in the atmospheric sciences. Academic press.
23. Hamilton, J.D., 2016. Macroeconomic Regimes and Regime Shifts. *Handbook of Macroeconomics - Chapter 3*, pp 163–201. Doi: [org/10.1016/bs.hesmac.2016.03.004](https://doi.org/10.1016/bs.hesmac.2016.03.004)
24. Zhou, Y., Wang, L., Zhong, R., and Tan, Y., 2018. A Markov Chain Based Demand Prediction Model for Stations in Bike Sharing Systems. *Mathematical Problems in Engineering*, 2018, pp.1–8. Doi: 10.1155/2018/8028714
25. Pinsky, M., and Karlin, S., 2010. An introduction to stochastic modeling. Academic press
26. Yousefi, H., Roumi, S., Tabasi, S., and Hamlehdar, M., 2018. Economic and air pollution effects of city council legislations on renewable energy utilisation in Tehran. *International Journal of Ambient Energy*, 39(6), pp.626–631. Doi: 10.1080/01430750.2017.1324819
27. Rahimi, J., Rahimi, A., and Bazrafshan, J., 2014. Study of persistence of polluted days with carbon monoxide (CO) in Tehran city using Markov Chain model. *Journal of Environmental Science and Technology*, 15(2), pp.79–90

#### COPYRIGHTS

©2024 The author(s). This is an open access article distributed under the terms of the Creative Commons Attribution (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, as long as the original authors and source are cited. No permission is required from the authors or the publishers.



#### Persian Abstract

#### چکیده

شاخص کیفیت هوا (AQI) عددی است که از آن برای توصیف وضعیت آلودگی هوا استفاده می‌شود و بوسیله آن کلاس شدت آلودگی هوا تعیین می‌گردد و نشان‌دهنده آن است درک رفتار تصادفی آلودگی هوا و تغییر طبقات شدت آلودگی هوا اهمیت زیادی در مدیریت بهداشت عمومی جامعه و پایداری شبکه بهداشت و درمان کشور دارد. در این راستا مدل زنجیره مارکف ابزاری مناسب برای تحلیل احتمالاتی انتقال بین طبقات شدت خشکسالی است. برای این منظور ضمن تعیین مرتبه مناسب زنجیره مارکف با شاخص BIC به تحلیل ماتریس‌های احتمال انتقال ۶ حالت آلودگی روزانه هوا شامل: پاک، سالم، ناسالم برای گروه‌های حساس، ناسالم، بسیار ناسالم و خطرناک در ۱۲ ایستگاه کیفیت سنج هوا در مناطق شمالی، جنوبی، شرقی، غربی و مرکزی شهر تهران طی سال‌های ۱۳۹۷-۱۴۰۱ پرداخته شد و پس از بررسی فراوانی کلاس‌های شدت آلودگی هوا به تحلیل ماتریس‌های احتمال انتقال و بررسی تداوم ثبات هر یک از کلاس‌ها پرداخته شد. نتایج نشان داد که مناطق غربی و جنوبی تهران آلوده‌ترین نقاط شهری می‌باشند به طوری که به ترتیب بیشتر از ۴۴ و ۳۴ درصد فراوانی مربوط به شاخص‌های ناسالم بوده است، همچنین ایستگاه تربیت مدرس در منطقه‌ی مرکزی بیش از ۴۰ درصد فراوانی شاخص ناسالم و ایستگاه پیروزی در منطقه شرقی ۳۹ درصد فراوانی شاخص ناسالم بوده‌اند و شمال تهران پاک‌ترین وضعیت هوا را دارد. نتایج ماتریس احتمال انتقال نشان داد که بیشترین تداوم ثبات برای همه مناطق مربوط به هوای سالم می‌باشد. همچنین احتمال انتقال هوای پاک و سالم به هوای بسیار ناسالم و خطرناک و بالعکس وجود ندارد و تداوم ثبات هوای خطرناک در اکثر ایستگاه‌ها به جز منطقه غربی و جنوبی (۲۵ درصد تداوم ثبات) صفر می‌باشد و احتمال تداوم ثبات هوای ناسالم در مناطق غربی و جنوبی ۶۰ درصد می‌باشد که نشان‌دهنده آلودگی بیشتر این مناطق نسبت به سایر مناطق دارد. نتایج ماتریس احتمال انتقال حاکی از آن بود که بیشترین احتمال استقرار آلودگی مربوط به مناطق غربی و جنوبی تهران بوده و پاک‌ترین هوا مربوط به ناحیه‌ی شمالی شهر تهران است به طوری احتمال تداوم هوای پاک در منطقه‌ی غربی کمتر از ۳۰ درصد و احتمال انتقال هوای بسیار ناسالم و خطرناک به هوای ناسالم ۵۰ درصد بوده است. اما احتمال تداوم ثبات شاخص هوای ناسالم در شمال تهران بیش از ۴۰ درصد می‌باشد و این احتمال برای هوای پاک و سالم بیشتر از ۵۰ درصد بوده است.